UNIVERSITAT DE BARCELONA
Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

The Radon Transform and Its Application to Computed Tomography

Autor: Agustín A. Sorrentino Rodríguez

Director: Dr. J. Abel Antezana
Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, June 10, 2026

Abstract

This work studies the mathematical foundations of computerized tomography from the point of view of the Radon transform. Starting from the physical interpretation of X-ray measurements as line integrals, we introduce the main analytical tools needed to formulate the reconstruction problem, including Fourier analysis, distributions, sampling theory and the Fourier slice theorem. We then discuss several inversion procedures, with special emphasis on filtered backprojection and Fourier-based reconstruction methods. Following the approach of Cascante et al., we also describe a gridding method for the numerical recovery of images from discrete projection data. The final part of the work analyzes the computational structure of the algorithm and the role of sampling, interpolation and filtering in practical reconstruction. Overall, the aim is to show how the continuous theory of the Radon transform leads naturally to implementable algorithms for medical image reconstruction.

2020 Mathematics Subject Classification. 44A12, 92C55, 65T50

Acknowledgements

En primer lugar no es solo un agradecimiento, sino algo más fuerte, para lo que no tengo palabras. Su confianza en mí, incluso en momentos en los que yo la perdía un poco en mí mismo, su optimismo, sus gestos de cariño, su practicidad e inteligencia son lo que me ha guiado durante estos años. Le agradezco a *Romi*, mi pareja, novia y madre de nuestra *Hija*.

En segundo lugar, mi *Mamá* ha jugado un rol de apoyo a pesar de la distancia, sin el cual dudo que hubiera podido mantenerme tan firme en este camino. Mis *Hermanos* y *Tito* siempre han estado presentes estos años, a pesar de la distancia. Agradezco también a mis *Tíos* de Barcelona.

Agradezco a los *profesores* de la facultad. *Jorge*, mi tutor de TFG, a quien "todavía" conozco poco (el "todavía" es aspiracional), por su gracia al explicar y por su cercanía desde el minuto cero. *Nahuel* Statuto y *Carles* Rovira no solo fueron buenos profesores, sino personas de confianza. Casi la totalidad de los profesores con los que tuve clases me inspiraron cosas buenas (no me perdí clases, incluso saliendo del trabajo, de C. Cascacuberta, T. Sanz, A. Viero, M. Lahoz, X. Massaneda, I. Mundet, J. Marzo, D. Marquez... y me faltan aún).

Agradezco a los compañeros (espero llamarlos "amigos" más tarde) que hice hace unos años (en particular a *A. Coll* y *A. Solé*) y a los que hice este último año.

Finalmente agradezco a mis *Compañeros de Neurología* de los hospitales *Germans Trias i Pujol* y *Moisés Broggi*, por su interés y apoyo continuo. Es un poco tonto, pero estoy agradecido con la medicina y la neurología, ya que me han permitido conocerme (y mantenerme económicamente), y me han mantenido mentalmente saludable estos años.

Contents

Abstract i

1 Introduction . . . 1
- 1.1 The reconstruction problem . . . 1
- 1.2 A first look at projections and the Radon transform . . . 3
- 1.3 A first look at reconstruction algorithms . . . 6
- 1.4 Structure of the work . . . 9

2 Preliminaries . . . 9

3 The Radon transform . . . 14
- 3.1 Definition and basic properties . . . 14
- 3.2 Uniqueness . . . 18
- 3.3 Range . . . 21
- 3.4 Inversion formulas . . . 25
  - 3.4.1 The Hilbert transform. . . . 27

4 Discretization of the Radon transform . . . 29
- 4.1 The semidiscrete case . . . 35

5 Reconstruction algorithms . . . 40
- 5.1 Filtered backprojection . . . 40
  - 5.1.1 Cost. . . . 44
  - 5.1.2 Algorithm. . . . 45
- 5.2 The gridding method . . . 45
  - 5.2.1 Discretization of bandlimited functions on $[0, \infty)$ . . . 46
  - 5.2.2 Algorithm . . . 49
  - 5.2.3 Cost. . . . 50

6 Conclusion . . . 50

Appendix A. The concentration operator . . . A-1
- A.1 Additional functional analysis preliminaries. A-1
- A.2 Proof of Prop. 4.3. . . . A-1

1 Introduction

1.1 The reconstruction problem

Tomographic imaging is based on an indirect observation principle: instead of measuring an object point by point, one records how it attenuates beams passing through it. The mathematical problem is then to determine whether, and how, the object can be reconstructed from these lower-dimensional measurements.

The following description follows [GW18, Sec. 5.11, pp. 368–397] and [Bea12]. We examine the problem of reconstructing an image from a series of projections, with a focus on *X*–ray *computed tomography*.

The reconstruction problem can be explained intuitively without equations. Consider Fig. 1(a): a single object on a uniform background. Suppose it is a *cross-section* of a 3-D region of a human head, that the background represents soft (uniform) tissue, and that the object is a tumor, also uniform, but with higher X-ray absorption characteristics.

Suppose next that we pass a thin, flat beam of X-rays from left to right through the plane of the image, as Fig. 1(b) shows, and assume that the energy of the beam is absorbed more by the object than by the background. Using a strip of X-ray absorption detectors on the other side of the region will yield the signal, or *absorption profile*, shown, whose amplitude is proportional to absorption. We may view any point in the signal as the sum of the absorption values across the single ray in the beam corresponding spatially to that point. Such a sum is referred to as a *raysum*. At this point, all the information we have about the object is this 1-D absorption signal.

We have no way of determining from a single projection whether we are dealing with a single object, or a multitude of objects along the path of the beam, but we begin the reconstruction by creating an image based only on this information. The approach is to project the 1-D signal *back* in the opposite direction from which the beam came, as Fig. 1(c) shows. The process of backprojecting a 1-D signal across a 2-D area is sometimes referred to as *smearing* the projection back across the area.

In terms of digital images, this means duplicating the same 1-D signal across the image, perpendicularly to the direction of the beam. For example, Fig. 1(c) was created by duplicating the 1-D signal in all columns of the reconstructed image. The approach just described is called *backprojection*.

Next, suppose that we rotate the position of the source-detector pair by 90°, as in Fig. 1(d). Repeating the procedure yields a backprojection image in the vertical direction (Fig. 1(e)). We continue the reconstruction by *adding* this result to the previous backprojection, resulting in Fig. 1(f). Now, we begin to suspect that the object of interest is contained in the square shown, whose amplitude is twice the

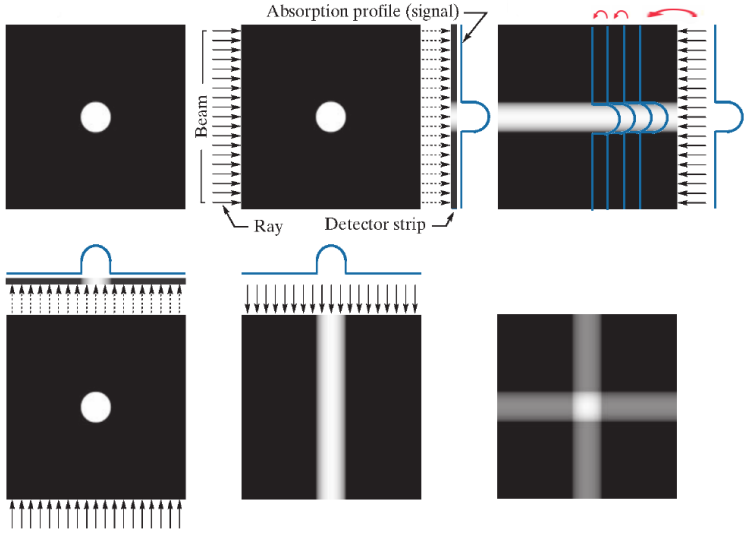


Figure 1: Panels are numbered from top to bottom and from left to right. **(a)** A planar region containing a single object. **(b)** A parallel beam, a detector array, and the measured one-dimensional absorption profile. **(c)** Result of backprojecting the absorption profile. **(d)** Beam and detectors rotated by 90°. **(e)** New backprojection. **(f)** Sum of both backprojections. Adapted from [GW18, Fig. 5.32].

amplitude of the individual backprojections because the signals were added. We should be able to learn more about the shape of the object in question by taking more views in the manner just described, as Fig. 2 shows. As the number of projections increases, the amplitude strength of non-intersecting backprojections decreases relative to the strength of regions in which multiple backprojections intersect. The net effect is that brighter regions will dominate the result, and backprojections with few or no intersections will fade into the background as the image is scaled for display.

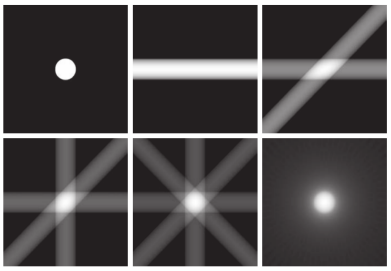


Figure 2: Panels are numbered from top to bottom and from left to right. **(a)** Same as Fig. 1(a). **(b)–(e)** Reconstruction using 1, 2, 3, and 4 backprojections 45° apart. **(f)** Reconstruction with 32 backprojections 5.625° apart. Notice the blurring. Adapted from [GW18, Fig. 5.33].

Fig. 2(f), which was formed from 32 backprojections, illustrates this concept. Note, however, that while this reconstructed image is a reasonably good approximation to the shape of the original object, the image is blurred by a "halo" effect, the formation of which can be seen in progressive stages in Fig. 2. We conclude from the discussion of Figs. 1 and 2 that backprojections 180° apart are mirror images of each other, so we have to consider only angle increments halfway around a circle in order to generate all the backprojections required for reconstruction.

Finally, Fig. 3 shows the same backprojection procedure applied to a planar region containing two objects with different absorption characteristics.

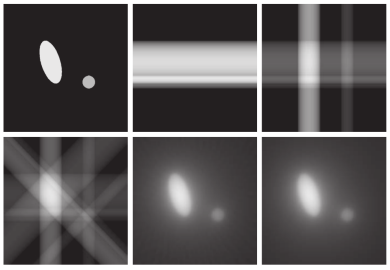
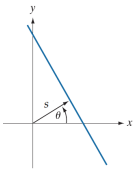


Figure 3: Panels are numbered from top to bottom and from left to right. **(a)** Two objects with different absorption characteristics. **(b)–(d)** Reconstruction using 1, 2, and 4 backprojections, 45° apart. **(e)** Reconstruction with 32 backprojections, 5.625° apart. **(f)** Reconstruction with 64 backprojections, 2.8125° apart. Adapted from [GW18, Fig. 5.34].

1.2 A first look at projections and the Radon transform[2]



A straight line in Cartesian coordinates can be described either by its *slope-intercept form*, $y = ax + b$, or by its *normal representation*:

$$L(\theta, s) := \left\{(x, y) \in \mathbb{R}^2 : x\cos\theta + y\sin\theta = s\right\}.$$

The projection of a *parallel-ray beam* can be modelled by a set of such lines, as Fig. 4 shows. A *point* (θ_k, s_j) in the projection profile is given by the *raysum* along the line $L(\theta_k, s_j)$. With continuous quantities, the *raysum* is a line integral, given by

$$g(\theta_k, s_j) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,\delta_0(x\cos\theta_k + y\sin\theta_k - s_j)\,dx\,dy,$$ [3]

If we consider all values of θ and s, the preceding equation generalizes to

$$g(\theta, s) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,\delta_0(x\cos\theta + y\sin\theta - s)\,dx\,dy. \tag{1}$$

This equation gives the projection (line integral) of $f(x,y)$ along a line in the xy-plane, and is known as the *Radon transform of* f. We willuse the notation $\mathcal{R}f$ instead of g. In the discrete case (1) becomes

$$\mathcal{R}f(\theta, s) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\,\delta_0(x\cos\theta + y\sin\theta - s), \tag{2}$$

where x, y, and s are now discrete variables, and M and N are the dimensions of the rectangular area over which the transform is applied.

If we fix θ and allow s to vary, then (2) simply sums values of $f(x,y)$ along $L(\theta, s)$ lines. Going through all values of s required to span the $M \times N$ area, yields *one* projection. Changing θ and repeating this procedure yields another projection, and so forth. This is how the projections in Figs. 1–3 were generated.

Sinograms. When $\mathcal{R}f(\theta, s)$ is displayed as an image with s and θ as rectangular coordinates, the result is called a *sinogram*, similar in concept to displaying

[2] For Fourier transform, inverse transform and convolution, see the Fourier part of the Preliminaries.

[3] $\delta_a(x) = 1$ if $a = x$ and $\delta_0(x) = 0$ if $x \neq a$.

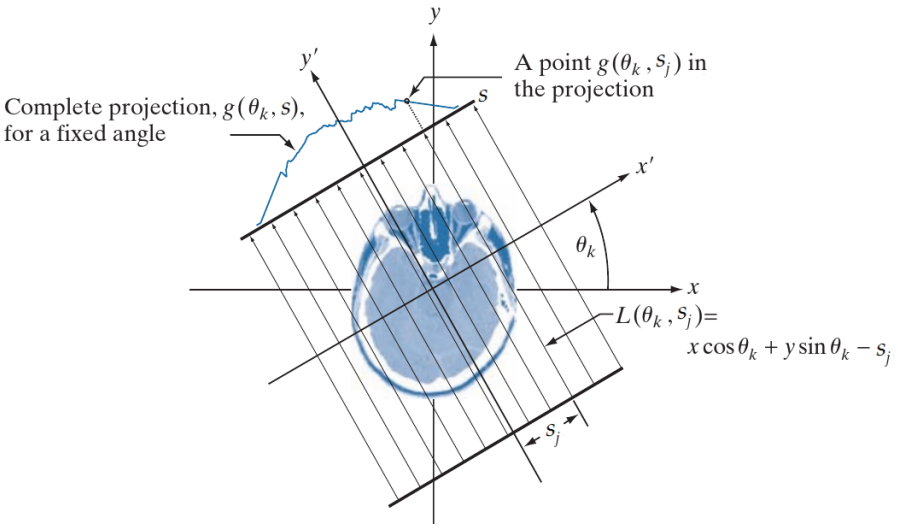


Figure 4: Geometry of a parallel-ray beam geometry. Adapted from [GW18, Fig. 5.37].

the Fourier spectrum. Like the Fourier transform, a sinogram contains the data necessary to reconstruct $f(x, y)$. Unlike the Fourier transform, $\mathcal{R}f(\theta, s)$ is always a real function whenever f is real-valued.

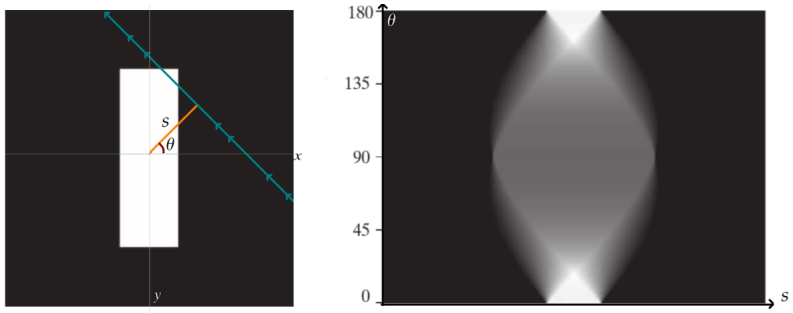


Figure 5: (a) A rectangular region. **(b)** Sinogram of the rectangle in **(a)**. The vertical and horizontal axes correspond to θ and s, respectively. The bottom row is the projection of the rectangle in the horizontal direction (i.e., $\theta = 0°$), and the middle row is the projection in the vertical direction ($\theta = 90°$). The fact that the nonzero portion of the bottom row is smaller than the nonzero portion of the middle row tells us that the object is narrower in the horizontal direction. The fact that the sinogram is symmetric in both directions about the center of the image tells us that we are dealing with an object that is symmetric and parallel to the x- and y-axes. Finally, the sinogram is smooth, indicating that the object has a uniform intensity. Adapted from [GW18, Fig. 5.39].

As is the case with displays of the Fourier spectrum, sinograms can be readily interpreted for but only for simple regions. For example, Fig. 5(b) is the sinogram of the rectangle on its left.

Backprojections. To obtain a formal expression for a backprojected image from $\mathcal{R}f$, let us begin with a *single* point $\mathcal{R}f(\theta_k, s_j)$ of the complete projection $\mathcal{R}f(\theta_k, s)$, for a fixed θ_k (see Fig. 4). Forming part of an image by backprojecting this single point is nothing more than copying the line $L(\theta_k, s_j)$ onto the image, where the value (intensity), of each point in that line is $\mathcal{R}f(\theta_k, s_j)$. Repeating this process for

all values of s_j in the projected signal, but keeping the value of θ fixed at θ_k, results in the following expression:

$$f_{\theta_k}(x,y) = \mathcal{R}f(\theta_k, s) = \mathcal{R}f\big(\theta_k, x\cos\theta_k + y\sin\theta_k\big)$$

for the image due to backprojecting the projection obtained with a fixed angle θ_k, as in Fig. 1(b). We may write, in general, that the image formed from a *single* backprojection obtained at an angle θ is given by

$$f_\theta(x,y) = \mathcal{R}f\big(\theta, x\cos\theta + y\sin\theta\big).$$

We form the final backprojected image by integrating over all the backprojected images:

$$f_{\text{bp}}(x,y) = \int_0^\pi f_\theta(x,y)\,d\theta. \tag{3}$$

For the rest of the Introduction, set $e_\theta := (\cos(\theta), \sin(\theta))$ and recall that $s = x\cos\theta + y\sin\theta$. We now introduce the notation $\mathcal{R}^\#$ for the *backprojection operator*

$$\mathcal{R}^\#(\mathcal{R}f)(x,y) := \int_0^{2\pi} \mathcal{R}f\big(\theta, x\cos\theta + y\sin\theta\big)\,d\theta = \int_0^{2\pi} \mathcal{R}f\big(\theta, (x,y)\cdot e_\theta\big)\,d\theta. \tag{4}$$

Notice that, however, that the angles θ and $\theta+\pi$ give the same projections, up to the change $s \mapsto -s$. Indeed, $\mathcal{R}f(\theta+\pi, -s) = \mathcal{R}f(\theta, s)$, as can be seen in Fig. 6.

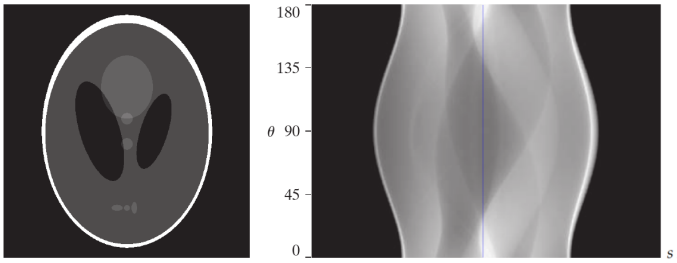


Figure 6: On the left, the Shepp–Logan phantom, a widely used synthetic image in medical image reconstruction. On the right is its sinogram, with s on the horizontal axis and θ on the vertical axis. Adapted from [GW18, Fig. 5.39].

Therefore, in (4), the integral over $[\pi, 2\pi]$ repeats the integral over $[0, \pi]$. Hence Eq. (3) can be rewritten as

$$f_{\text{bp}}(x,y) = \int_0^\pi f_\theta(x,y)\,d\theta = \int_0^\pi \mathcal{R}f\big(\theta, x\cos\theta + y\sin\theta\big)\,d\theta = \frac{1}{2}\,\mathcal{R}^\#(\mathcal{R}f)(x,y).$$

Using the discrete version of Eq. (3), we obtain the backprojected image shown in Fig. 7. We observe *blurring*: the rectangle is located correctly, but is surrounded by a diffuse halo. A backprojected image formed in the manner just described sometimes is referred to as a *laminogram*. Such an image is only an approximation to the original one.



Figure 7: Backprojection of the sinogram in Fig. 5.

The Fourier projection–slice theorem. This theorem gives a more precise relation between the projection data and the original image, useful to remove the blurring. For a fixed angle θ, take the 1-D Fourier transform of $\mathcal{R}f(\theta,\cdot)$ (i.e. the function $s \mapsto \mathcal{R}f(\theta,s)$). From now on

$$\widehat{\mathcal{R}f}(\theta,\sigma) := \mathcal{F}_s(\mathcal{R}f)(\theta,\sigma) = \int_{-\infty}^{\infty} \mathcal{R}f(\theta,s)\, e^{-2\pi i\sigma s}\, ds.$$

On the other hand, the original image $f(x,y)$ has a 2-D Fourier transform,

$$\mathcal{F}_{x,y}f(u,v) = \widehat{f}(u,v) = \int_{\mathbb{R}^2} f(x,y)\, e^{-2\pi i(ux+vy)}\, dx\, dy.$$

The Fourier–slice theorem states that the one-dimensional Fourier transform of a projection gives the values of $\widehat{f}(u,v)$ on a line in the Fourier plane. More precisely, the change of variables $(u,v) = (\sigma\cos\theta, \sigma\sin\theta)$ for $\sigma \in \mathbb{R} = \sigma e_\theta$, gives

$$\widehat{\mathcal{R}f}(\theta,\sigma) = \widehat{f}(u,v) = \widehat{f}(\sigma\cos\theta, \sigma\sin\theta). \tag{5}$$

Notice that this same vector e_θ is normal to the spatial lines $L(\theta,s)$, but it is the direction vector of the corresponding line in the Fourier plane. This correspondence is illustrated in Fig. 8.

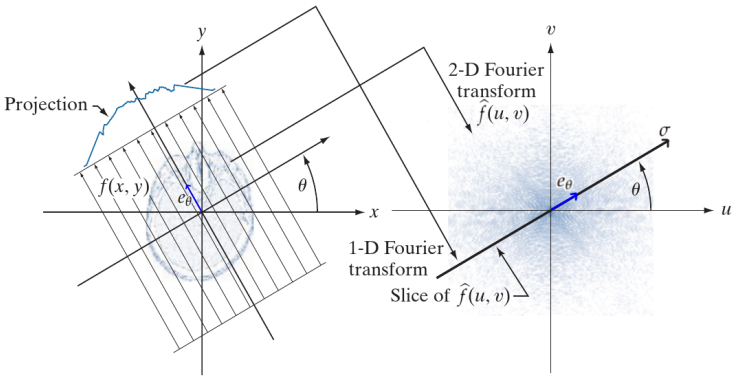


Figure 8: The Fourier projection–slice theorem. $\widehat{\mathcal{R}f}(\theta,\sigma)$ is equal to $\widehat{f}(u,v)$ along the radial line $(u,v) = \sigma(\cos\theta, \sin\theta) = \sigma e_\theta$. Thus, the angle used to generate the projection is the same angle that determines the corresponding slice in the Fourier domain. Adapted from [GW18, Fig. 5.41].

If we knew $\mathcal{R}f(\theta,s)$ for all angles θ and all positions s, we would know the values of $\widehat{f}(u,v)$ along all radial lines in the Fourier plane and then recover $f(x,y)$ by inverse 2-D Fourier transform. However, this is computationally expensive.

1.3 A first look at reconstruction algorithms

Projection–slice leads to two reconstruction routes: "filtered backprojection" in physical space and "gridding" in frequency space. Here we only sketch the first.

Filtered backprojection. Recalling that $xu + yv = (x,y)\cdot(u,v) = (x,y)\cdot\sigma e_\theta = \sigma\,((x,y)\cdot e_\theta)$, we obtain

$$f(x,y) = \int_{\mathbb{R}^2} \widehat{f}(u,v) e^{2\pi i(ux+vy)}\,du\,dv \overset{du\,dv=|\sigma|\,d\sigma\,d\theta}{=} \frac{1}{2}\int_0^{2\pi}\int_{-\infty}^{\infty} \widehat{f}(\sigma e_\theta) e^{2\pi i\sigma((x,y)\cdot e_\theta)}|\sigma|\,d\sigma\,d\theta$$
$$\overset{\text{Eq. (5)}}{=} \frac{1}{2}\int_0^{2\pi} \underbrace{\int_{-\infty}^{\infty} \widehat{\mathcal{R}f}(\theta,\sigma)\, e^{2\pi i\sigma((x,y)\cdot e_\theta)}|\sigma|\,d\sigma}_{I(\theta)}\,d\theta. \tag{6}$$

The factor $1/2$ appears because the parametrization $(u,v) = \sigma e_\theta$, with $\sigma \in \mathbb{R}$ and $\theta \in [0, 2\pi]$, covers each point of $\mathbb{R}^2 \setminus \{0\}$ twice.

Notice that the factor $|\sigma|$ pops up. Without it, the inner integral would simply recover the projection $\mathcal{R}f(\theta,s)$, and the result would be the *unfiltered* backprojection. Thus, the factor $|\sigma|$ is precisely what corrects the blurring. However, it is not useful in practice: $|\sigma|$ grows without bound and in particular, it is not integrable on $\mathbb{R}$, so it cannot be used in computations.[4] To deal with this, since $(x,y)\cdot e_\theta = x\cos\theta + y\sin\theta = s$, we rewrite the inner integral $I(\theta)$ in (6):

$$I(\theta) = \int_{-\infty}^{\infty} |\sigma|\,\widehat{\mathcal{R}f}(\theta,\sigma)\, e^{2\pi i\sigma s}\,d\sigma\Big|_{s=(x,y)\cdot e_\theta} = \mathcal{F}_\sigma^{-1}\left(|\sigma|\,\widehat{\mathcal{R}f}(\theta,\sigma)\right)(s)\Big|_{s=(x,y)\cdot e_\theta}.$$

Substituting this expression into Eq. (6), we obtain

$$2f(x,y) = \int_0^{2\pi} \mathcal{F}_\sigma^{-1}\left(|\sigma|\,\widehat{\mathcal{R}f}(\theta,\sigma)\right)(s)\Big|_{s=(x,y)\cdot e_\theta} d\theta = \mathcal{R}^{\#}\mathcal{F}_\sigma^{-1}\left(|\sigma|\,\widehat{\mathcal{R}f}(\theta,\sigma)\right)(x,y), \tag{7}$$

where the last equality follows from Eq. (4) applied to $\mathcal{F}_\sigma^{-1}\left(|\sigma|\,\widehat{\mathcal{R}f}(\theta,\sigma)\right)(s)$

Although Eq. (7) gives the correct inversion formula, it is not yet written in the most useful way. To interpret $\mathcal{F}_\sigma^{-1}\left(|\sigma|\,\widehat{\mathcal{R}f}(\theta,\sigma)\right)(s)$ suppose that there exists a function $\varphi\colon \mathbb{R}\to\mathbb{C}$ such that $\widehat{\varphi}(\sigma) = |\sigma|$. Then

$$|\sigma|\,\widehat{\mathcal{R}f}(\theta,\sigma) = \widehat{\varphi}(\sigma)\,\widehat{\mathcal{R}f}(\theta,\sigma) = \widehat{(\varphi *_s \mathcal{R}f)}(\theta,\sigma),$$

where $(\varphi *_s \mathcal{R}f)(\theta,s) := \int_{\mathbb{R}} \varphi(s-t)\,\mathcal{R}f(\theta,t)\,dt$. Therefore

$$\mathcal{F}_\sigma^{-1}\left(|\sigma|\,\widehat{\mathcal{R}f}(\theta,\sigma)\right)(s) = (\varphi *_s \mathcal{R}f)(\theta,s).$$

Substituting this expression into Eq. (7), we obtain the formal identity

$$f(x,y) = \frac{1}{2}\,\mathcal{R}^{\#}(\varphi *_s \mathcal{R}f)(x,y).$$

[4] From a practical point of view, using this exact multiplier would amplify arbitrarily high frequencies, precisely where measured and sampled projection data are most sensitive to noise.

Note that before applying $\mathcal{R}^{\#}$, each projection $s \mapsto \mathcal{R}f(\theta, s)$ is filtered by convolution with the same one-dimensional function φ. In other words, the reconstruction is obtained by backprojecting filtered projections.

There is, however, an obstruction. The identity $\widehat{\varphi}(\sigma) = |\sigma|$ cannot hold for an ordinary integrable function φ. Indeed, if $\varphi \in L^1(\mathbb{R})$, then the Riemann–Lebesgue lemma[5] gives $\widehat{\varphi}(\sigma) \longrightarrow 0$ as $|\sigma| \to \infty$. But $|\sigma| \longrightarrow \infty$ as $|\sigma| \to \infty$. Thus the ideal multiplier $|\sigma|$ does not come from an ordinary integrable function.

In practice, one replaces the ideal multiplier $|\sigma|$ by a multiplier with compact support [GW18, Sec. 5.11, pp. 381–388]. The simplest approach is to take $m_L(\sigma) = 0$ if $\sigma > L$, for $L > 0$. Since m_L vanishes outside a bounded interval, the filtered projection $\mathcal{F}^{-1}\left(m_L(\sigma)\,\widehat{\mathcal{R}f}(\theta,\sigma)\right)(s)$ is band-limited in the s-variable. With this replacement, Eq. (7) becomes the practical filtered backprojection formula

$$f(x,y) \approx \frac{1}{2}\,\mathcal{R}^{\#}\left[\mathcal{F}^{-1}\left(m_L(\sigma)\,\widehat{\mathcal{R}f}(\theta,\sigma)\right)\right](x,y).$$

If $m_L = \widehat{\phi_L}$, then the same convolution argument gives the equivalent form

$$f(x,y) \approx \frac{1}{2}\,\mathcal{R}^{\#}(\phi_L *_s \mathcal{R}f)(x,y).$$

The parameter L controls the frequencies used in the reconstruction. A larger value of L preserves more high-frequency information, but also amplifies more noise. A smaller value of L suppresses more high frequencies, but produces a smoother reconstruction. The choice of m_L is empirical. [6]

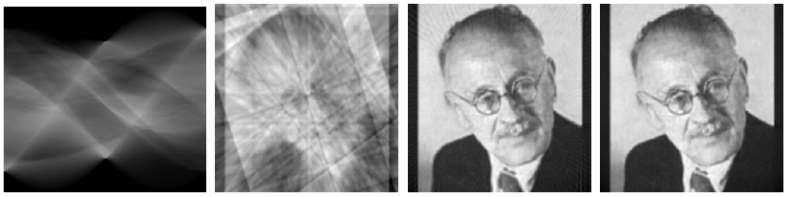


Figure 9: Sinogram and filtered backprojections using a Hann filter. From left to right: sinogram, reconstruction with 10 backprojections, reconstruction with 100 backprojections, and reconstruction with 1000 backprojections. Adapted from [Bea12, Fig. 8].

[5] If $h \in L^1(\mathbb{R})$, then $\widehat{h} \in C_0(\mathbb{R})$.

[6] We mention two standard choices of m_L. The *Ram–Lak filter* is the truncated ramp $m_L^{\mathrm{RL}}(\sigma) = |\sigma|\,\chi_L(\sigma)$. A smoother alternative is the *Hann filter*, defined by

$$m_L^{\mathrm{Hann}}(\sigma) = |\sigma|\,\frac{1}{2}\left(1 + \cos\frac{\pi\sigma}{L}\right)\chi_L(\sigma).$$

Compared with the Ram–Lak filter, the Hann filter attenuates high frequencies more gradually near the cut-off frequency. This reduces the amplification of high-frequency noise, at the cost of producing a slightly smoother reconstruction [Bea12].

1.4 Structure of the work

This work follows the general approach of Cascante [CMOCP07]. Section 2 collects the analytic preliminaries needed later, including material not covered in the compulsory courses of the degree. Section 3 studies the Radon transform and its structural properties. Section 4 discusses what can be recovered from finitely many projections. Section 5 presents the corresponding reconstruction algorithms.

2 Preliminaries

Throughout, all functions are complex-valued unless otherwise stated, and all Euclidean measures are Lebesgue measures. If the integration is over a k-dimensional affine subspace, dx denotes the corresponding k-dimensional Lebesgue measure. Lebesgue measure is *invariant under rigid motions*; in particular, if $Q \in O(n)$ and $a \in \mathbb{R}^n$, then $|QA + a| = |A|$ and $\int_{QA+a} f(x)\,dx = \int_A f(Qy + a)\,dy$. [Cer10, p. 10].

Throughout the rest of the work, for $n \geq 2$ we write

$$Z := S^{n-1} \times \mathbb{R}, \qquad \theta_s := \{x \in \mathbb{R}^n : x \cdot \theta = s\}.$$

The parameter θ is the *unit normal direction* and s is the *signed distance* from the origin. Thus $\theta_s = s\theta + \theta^\perp$, where $\theta^\perp := \{y \in \mathbb{R}^n : y \cdot \theta = 0\}$. Notice that $\theta_s = (-\theta)_{-s}$, so the same hyperplane has two representatives in Z.

The following preliminaries are meant to make the work self-contained for readers with the background provided by the compulsory courses of the Mathematics degree at our University. We collect the main conventions and standard formulas used later; some may be skipped on a first reading and recalled when needed.

Graph and polar integration. We shall use the formula for the *surface integral*: if $\Psi : U \subset \mathbb{R}^{n-1} \to M \subset \mathbb{R}^n$ parametrizes the hypersurface M, then

$$\int_M f\,d\sigma = \int_U f(\Psi(u))\sqrt{\det G(u)}\,du, \qquad G_{ij}(u) := \partial_i\Psi(u) \cdot \partial_j\Psi(u).$$

Equivalently, if $N(u)$ is a unit normal,

$$\sqrt{\det G(u)} = \left|\det\big(N(u), \partial_1\Psi(u), \ldots, \partial_{n-1}\Psi(u)\big)\right|,$$

[HH15, Def. 5.3.2, Eq. (5.3.5), p. 532; Prop. 5.1.1, pp. 528–529].

For a graph $\Psi(u) = (u, \varphi(u))$, since $G = I + \nabla\varphi(\nabla\varphi)^t$, we have

$$\int_{\Gamma_\varphi} f\,d\sigma = \int_U f(u, \varphi(u))\sqrt{1 + |\nabla\varphi(u)|^2}\,du,$$

We also use *polar coordinates*:

$$\int_{\mathbb{R}^n} f(x)\,dx = \int_0^\infty \int_{S^{n-1}} f(r\theta) r^{n-1}\,d\theta\,dr,$$

where $d\theta$ denotes surface measure on S^{n-1} [Sil26, Ex. 0.5, pp. 11–12].

Growth. For functions $u, v \geq 0$, $u = O(v)$ as $|x| \to \infty$ means that $u(x) \leq Cv(x)$ outside a large ball. If $f \in C(\mathbb{R}^n)$ and $|f(x)| = O(|x|^{-a})$, we shall also write, after enlarging C, $|f(x)| \leq C(1+|x|)^{-a}$ for all x. We shall use

$$(1+|\cdot|)^{-N} \in L^p(\mathbb{R}^k) \quad \Longleftrightarrow \quad Np > k, \qquad k \geq 1.^{7} \tag{8}$$

A function f is *rapidly decreasing* on $\mathbb{R}^n$ if, for every N, $|f(x)| \leq C_N(1+|x|)^{-N}$ for all x; on Z this means $|g(\theta,s)| \leq C_N(1+|s|)^{-N}$ with C_N independent of θ. We use *polynomial growth* (or *slow increase*) when $|f(x)| = O((1+|x|)^N)$ for some N.

Continuity of parametric integrals. If $F(p,\cdot)$ is measurable, $F(\cdot,x)$ is a.e. continuous and locally L^1-dominated, then $p \mapsto \int_X F(p,x)\,d\mu(x)$ is continuous, by dominated convergence [Cer10, p. 12].

Test and Schwartz functions. C_c^∞ denotes smooth compactly supported functions; on $Z = S^{n-1} \times \mathbb{R}$, compact support is understood in the s-variable. They are usually called *test functions*. The *Schwartz space* on $\mathbb{R}^n$ is

$$\mathcal{S}(\mathbb{R}^n) = \left\{ f \in C^\infty : \sup_{x \in \mathbb{R}^n} (1+|x|)^N |\partial^\alpha f(x)| < \infty, \quad \forall N, \alpha \right\},$$

and define $\mathcal{S}(Z)$ similarly, using $D_\theta^\alpha \partial_s^\beta$ and decay in s. Thus Schwartz functions and all their derivatives decay rapidly; in particular the polynomially weighted products used below are absolutely integrable, so Tonelli–Fubini applies. Finally,

$$C_c^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n),$$

with both inclusions strict.[8]

L^2-spaces and operators. For a measure space (X,μ), $L^2(X,\mu)$ denotes the Hilbert space of square-integrable functions modulo a.e. equality, with scalar product $\langle f, g\rangle_{L^2(X,\mu)} := \int_X f(x)\overline{g(x)}\,d\mu(x)$, which is linear in the first variable and conjugate-linear in the second [Cer10, Ex. 2.33, p. 47]. Unless otherwise stated, subsets of $\mathbb{R}^n$ carry Lebesgue measure, S^{n-1} carries surface measure $d\theta$, and $Z = S^{n-1} \times \mathbb{R}$ carries $d\theta\,ds$.

A Hilbert space is *separable* if it has a countable dense subset [RS80, p. 47]. An *orthonormal basis* is a complete orthonormal family, i.e. $f \perp e_k$ for all k implies

[7] Indeed, $\|(1+|\cdot|)^{-N}\|_p^p = \int_{\mathbb{R}^k} (1+|x|)^{-Np}\,dx \simeq 1 + \int_1^\infty r^{k-1-Np}\,dr$, which is finite iff $Np > k$. See [HH15, Sec. 4.10].

[8] The first inclusion is immediate. If $f \in \mathcal{S}$, then $|f(x)| \leq C_N(1+|x|)^{-N}$, hence $f \in L^1 \cap L^2$ by (8). Strictness follows from $e^{-|x|^2} \in \mathcal{S} \setminus C_c^\infty$ and $(1+|x|^2)^{-(n+1)/2} \in L^1 \cap L^2 \setminus \mathcal{S}$.

$f = 0$, equivalently $\overline{\text{span}}\{e_k\} = H$ [Cer10, p. 50]. In a separable Hilbert space, orthonormal bases are finite or countable [Cer10, p. 50]; hence, for such a basis,

$$f = \sum_k \langle f, e_k \rangle \, e_k, \qquad \|f\|^2 = \sum_k |\langle f, e_k \rangle|^2$$

[Cer10, Thm. 2.37, p. 50]. The spaces $L^2(\mathbb{R}^n)$ and $L^2(D)$, for a disc or ball $D \subset \mathbb{R}^n$, are separable [Grö01, Ch. 5, p. 83] [ABR01, Prop. 8.4, proof, p. 175]. For a closed subspace $V \subset L^2(X, \mu)$, $\dim V$ is the cardinality of any orthonormal basis of V. The *orthogonal projection* P_V onto V is defined by $P_V f \in V$ and $f - P_V f \perp V$; moreover

$$d(f, V) = \inf_{v \in V} \|f - v\|_2 = \|f - P_V f\|_2 .$$

Below, H denotes a Hilbert space, usually $L^2(\mathbb{R}^n)$ or a closed subspace of it. A set is relatively compact if its closure is compact, and a linear operator $T : H \to H$ is *compact* if it maps bounded sets into relatively compact sets [RS80, Ch. VI, Sec. 5, p. 199]. A linear operator $T : H \to H$ is *bounded* if $\|Tf\| \leq C \|f\|$ [Cer10, Thm. 2.21, p. 39], *self-adjoint* if $\langle Tf, g \rangle = \langle f, Tg \rangle$ [Cer10, Sec. 4.1, p. 97], *positive* if $\langle Tf, f \rangle \geq 0$ [Cer10, Thm. 8.25, p. 249] and *unitary* if it is bijective and $\langle Tf, Tg \rangle = \langle f, g \rangle$ [Cer10, Sec. 8.5.3, p. 248]. An *eigenvalue* is a $\lambda \in \mathbb{C}$ such that $Tf = \lambda f$ for some $f \neq 0$ [Cer10, Sec. 4.4.2, p. 118]. If T is self-adjoint, then $\lambda \|f\|^2 = \langle Tf, f \rangle = \langle f, Tf \rangle = \overline{\lambda} \|f\|^2$, so $\lambda \in \mathbb{R}$; if also $T \geq 0$, then $\lambda \|f\|^2 = \langle Tf, f \rangle \geq 0$. The *spectral theorem* [Cer10, Thm. 4.36, p. 121] says that, if T is compact, self-adjoint and positive on a separable Hilbert space H, then H admits an orthonormal basis of eigenfunctions (e_k), with $Te_k = \lambda_k e_k$, $\lambda_k \geq 0$, and the non-zero eigenvalues can accumulate only at 0. If, moreover, T is injective, then $\lambda_k > 0$, and

$$Tf = \sum_k \lambda_k \langle f, e_k \rangle \, e_k, \qquad f \in H.$$

Thus the relevant finite-dimensional spaces below will often be written as $\text{span}\{e_1, \ldots, e_N\}$.

P_V is linear, self-adjoint, $P_V^2 = P_V$, and $\|P_V\| \leq 1$ [Cer10, Thm. 2.35, pp. 47–48].

Fourier transform. For $f \in L^1(\mathbb{R}^n)$, its *Fourier transform* (FT) is defined, as in [Cer10, Sec. 7.1, p. 182], by

$$\widehat{f}(\xi) = \mathcal{F} f(\xi) := \int_{\mathbb{R}^n} f(x) e^{-2\pi i x \cdot \xi} \, dx.$$

FT is injective on $L^1(\mathbb{R}^n)$, as a consequence of the inversion formula [Grö01, Thm. 1.2.2, p. 10]. When restricted to $\mathcal{S}(\mathbb{R}^n)$, it is a linear homeomorphism $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$, with inverse [Cer10, Thm. 7.5, pp. 188–189]

$$\check{h}(x) = \mathcal{F}^{-1} h(x) = \int_{\mathbb{R}^n} h(\xi) e^{2\pi i x \cdot \xi} \, d\xi.$$

The same notation will be used on Z, but always in the real variable s:

$$\widehat{g}(\theta,\sigma) = \mathcal{F}_s(g(\theta,\cdot))(\sigma) := \int_{\mathbb{R}} g(\theta,s)e^{-2\pi i s\sigma}\,ds,$$

unless another variable is explicitly indicated. For $f \in \mathcal{S}(\mathbb{R}^n)$, [Cer10, Sec. 7.1, p. 182]

$$\widehat{\partial^\alpha f}(\xi) = (2\pi i\xi)^\alpha \widehat{f}(\xi), \qquad \widehat{x_j f}(\xi) = -\frac{1}{2\pi i}\partial_{\xi_j}\widehat{f}(\xi).$$

If $u, v \in L^1(\mathbb{R}^n)$, write [Grö01, Sec. 1.2, Eqs. (1.11)–(1.13), pp. 8–9]

$$(u * v)(x) := \int_{\mathbb{R}^n} u(x-y)v(y)\,dy; \qquad \widehat{u * v} = \widehat{u}\,\widehat{v}.$$

In the s-variable on Z, $(g *_s h)(\theta,s) := \int_{\mathbb{R}} g(\theta,s-t)h(\theta,t)\,dt$.

Applying the FT on $f \in \mathcal{S}(\mathbb{R}^n)$ twice gives reflection $\mathcal{F}^2 f(x) = f(-x)$, and $\check{f}(x) = \widehat{f}(-x)$ [Cer10, Thm. 7.5, pp. 188–189]. The same holds for $\mathcal{F}_s$ in the real variable s for $g \in \mathcal{S}(Z)$, $\mathcal{F}_s^2 g(\theta,s) = g(\theta,-s)$. Moreover, the FT preserves parity; if $u \in L^1(\mathbb{R}^n)$ is real-valued, then $\widehat{u}(-\xi) = \overline{\widehat{u}(\xi)}$. Thus real even functions have real even FTs [Grö01, Sec. 1.2, Eqs. (1.14)–(1.16), pp. 8–9].

Plancherel, Parseval and *Direct FT* (DFT). By Plancherel's theorem, $\mathcal{F}$ extends uniquely to a unitary operator on $L^2(\mathbb{R}^n)$ [Grö01, Thm. 1.1.2, p. 5]; hence, for $f, g \in L^2(\mathbb{R}^n)$,

$$\left\|\widehat{f}\right\|_2 = \|f\|_2, \qquad \left\langle \widehat{f}, \widehat{g}\right\rangle_2 = \langle f, g\rangle_2$$

the second identity being Parseval's identity [Grö01, Thm. 1.1.2, p. 5]. The finite analogue is the DFT on $G = \mathbb{Z}_M \times \mathbb{Z}_N$ [GW18, Sec. 4.5.5, Eqs. (4-67)–(4-68)]. For an $M \times N$ array $f = (f_{m,n})$, where $f_{m,n}$ denotes the value of f at the Cartesian point (m,n). Its 2-D DFT, and its 2-D inverse FT are, respectively

$$\widehat{f}_{k,\ell} = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f_{m,n}e^{-2\pi i(km/M+\ell n/N)}, \qquad f_{m,n} = \sum_{k=0}^{M-1}\sum_{\ell=0}^{N-1} \widehat{f}_{k,\ell}e^{2\pi i(km/M+\ell n/N)}.$$

where $\widehat{f}_{k,\ell}$ is the value of $\widehat{f}$ at the Fourier–variable point (k,l). Direct computation of the 2-D DFT costs $O(M^2N^2)$, while the *Fast FT* (FFT) reduces the corresponding cost to $O(MN\log(MN))$ [GW18, Sec. 4.11, Eq. (4-173)].

Tempered Distributions. [Cer10, Secs. 7.2–7.4, pp. 186–199]. We write $\mathcal{S}'(\mathbb{R}^n) := \mathcal{S}(\mathbb{R}^n)'$ for the space of *tempered distributions*, i.e. continuous linear functionals on $\mathcal{S}(\mathbb{R}^n)$. The action of $T \in \mathcal{S}'(\mathbb{R}^n)$ on $\varphi \in \mathcal{S}(\mathbb{R}^n)$ is denoted by $\langle T, \varphi\rangle$. If $1 \le p \le \infty$ and $(1+|x|)^{-N} f \in L^p(\mathbb{R}^n)$ for some N, then f defines a tempered distribution by [Cer10, Sec. 7.3, Ex. 7.8, p. 190]

$$\langle T_f, \varphi\rangle := \int_{\mathbb{R}^n} f(x)\varphi(x)\,dx.^{9}$$

[9] $$\int_{\mathbb{R}^n} |f(x)\varphi(x)|\,dx = \int_{\mathbb{R}^n} |(1+|x|)^{-N}f(x)|\,|(1+|x|)^N\varphi(x)|\,dx \overset{\text{Hölder}}{\le} \left\|(1+|\cdot|)^{-N}f\right\|_p \left\|(1+|\cdot|)^N\varphi\right\|_{p'},$$

We call T_f the *regular tempered distribution* induced by f. When a function induces a regular distribution, we often identify it with that distribution and simply write f instead of T_f; for instance, 1 may denote both the constant function and the tempered distribution $\varphi \mapsto \int_{\mathbb{R}^n} \varphi(x)\,dx$.

Convergence in $\mathcal{S}'$ means that $T_j \to T$ if $\langle T_j, \varphi\rangle \to \langle T, \varphi\rangle$ for $\varphi \in \mathcal{S}(\mathbb{R}^n)$. $T, S \in \mathcal{S}'$ are equal (denoted by $T \overset{\mathcal{S}'}{=} S$) if and only if $\langle T, \varphi\rangle = \langle S, \varphi\rangle$ for every $\varphi \in \mathcal{S}(\mathbb{R}^n)$. If two continuous functions f, g are equal in this sense, then they are equal pointwise.[10]

If $T \in \mathcal{S}'(\mathbb{R}^n)$ and $\psi \in \mathcal{S}(\mathbb{R}^n)$, define $(T * \psi)(x) := \langle T, \psi(x - \cdot)\rangle$. For $T = T_f$ this is the usual *convolution*:

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

The *FT* of $T \in \mathcal{S}'(\mathbb{R}^n)$ is defined by $\langle \widehat{T}, \varphi\rangle := \langle T, \widehat{\varphi}\rangle$, $\varphi \in \mathcal{S}(\mathbb{R}^n)$. If $f \in L^1(\mathbb{R}^n)$, then $\widehat{T_f} = T_{\widehat{f}}$, since

$$\langle \widehat{T_f}, \varphi\rangle = \int f(x)\widehat{\varphi}(x)\,dx = \int \widehat{f}(\xi)\varphi(\xi)\,d\xi.$$

Whenever the terms are defined, $\widehat{T * \psi} = \widehat{T}\,\widehat{\psi}$.

If $m \in O_M(\mathbb{R}^n)$, i.e. $m \in C^\infty$ and all its derivatives have at most polynomial growth, then *multiplication by m* maps $\mathcal{S}(\mathbb{R}^n)$ into itself. Thus, for $T \in \mathcal{S}'(\mathbb{R}^n)$, one defines

$$\langle mT, \varphi\rangle := \langle T, m\varphi\rangle, \qquad \varphi \in \mathcal{S}(\mathbb{R}^n), \quad \text{(see [RS80, Sec. V.3, p. 137])}.$$

The *Dirac mass* is the tempered distribution $\langle \delta_0, \varphi\rangle := \varphi(0)$, and $\widehat{\delta_0} = 1$.

Paley–Wiener spaces and sampling theorem. For a bounded measurable set $E \subset \mathbb{R}^n$, we set

$$PW_E := \left\{ f \in L^2(\mathbb{R}^n) : \operatorname{supp} \widehat{f} \subset \overline{E} \right\},$$

where $\widehat{f}$ is the L^2-Fourier transform defined by Plancherel's extension [Cer10, Thm. 7.13, pp. 192–193]. Elements of PW_E are called *E-bandlimited functions*. Since E is bounded, each class in PW_E has a canonical entire representative, given by $f(z) = \int_{\overline{E}} \widehat{f}(\xi) e^{2\pi i z\cdot\xi}\,d\xi$, $z \in \mathbb{C}^n$, which agrees with f a.e. on $\mathbb{R}^n$.

We will use Shannon's theorem as in [Nat01, Thm. 4.2]. Let $\Omega > 0$, and let $u, w \in L^2(\mathbb{R})$ be Ω–bandlimited. If $s_k = k\Delta s$ and $0 < \Delta s \le \frac{1}{2\Omega}$, then

$$\int_{\mathbb{R}} u(t)\overline{w(t)}\,dt = \Delta s \sum_{k\in\mathbb{Z}} u(s_k)\overline{w(s_k)}.$$

where $1 \le p \le \infty$ and p' is the conjugate exponent. Hence the integral is finite and depends continuously on $\varphi \in \mathcal{S}(\mathbb{R}^n)$, since $\varphi \in \mathcal{S}(\mathbb{R}^n)$, the function $(1 + |\cdot|)^N \varphi$ belongs to $L^{p'}(\mathbb{R}^n)$, because Schwartz functions decay faster than any power.

[10] $\int_{\mathbb{R}^n} (f - g)\varphi\,dx = 0\ \forall \varphi \in \mathcal{S}(\mathbb{R}^n) \Rightarrow f = g$ a.e.; by continuity, $f = g$.

3 The Radon transform

3.1 Definition and basic properties

Definition 3.1. *Let $f : \mathbb{R}^n \to \mathbb{C}$ be integrable on every hyperplane. The **Radon transform** of f is the function $\mathcal{R}f : Z \to \mathbb{C}$ defined by*

$$\mathcal{R}f(\theta, s) = \mathcal{R}_\theta f(s) := \int_{\theta_s} f(x)\, dx = \int_{\theta^\perp} f(s\theta + y)\, dy, \qquad (\theta, s) \in Z.$$

In particular, if $f \in C(\mathbb{R}^n)$ and $|f(x)| = O(|x|^{-a})$ for some $a > n-1$, $\mathcal{R}f(\theta, s)$ is well defined.[11]

If f is assumed to be real-valued, its Radon transform $\mathcal{R}f$ is also real-valued. For a real valued function $f : \mathbb{R}^3 \to \mathbb{R}$, picture a plane with unit normal vector $\theta \in S^2$ and signed distance s from the origin (recall representations are not unique). Now fix θ (i.e. fix a direction and thus a family of parallel planes). This gives the function $\mathcal{R}f(\theta, \cdot) : \mathbb{R} \to \mathbb{R}$. Now with that θ, vary s (i.e. vary the plane–origin distance). Finally, stop at some $\widetilde{s}$ and calculate $\mathcal{R}f(\theta, \widetilde{s})$, i.e. integrate f on the plane $\theta_{\widetilde{s}}$ (this will demand to parameterize such plane).

Proposition 3.2. ***Elementary symmetries.*** *Let f be integrable on every hyperplane. Then:*

(i) If $a \in \mathbb{R}^n$ and $(\tau_a f)(x) := f(x - a)$, then $\mathcal{R}(\tau_a f)(\theta, s) = \mathcal{R}f(\theta, s - \theta \cdot a)$.

(ii) If $U : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal and $g(x) := f(Ux)$, then $\mathcal{R}g(\theta, s) = \mathcal{R}f(U\theta, s)$.

These can be shown using change of variables ($u = x - a$ and $y = Ux$). One can choose an orthogonal U with $U\theta^* = \theta$, where $\theta^* = (0, \dots, 0, 1)$. Then $\mathcal{R}(f \circ U)(\theta^*, s) = \mathcal{R}f(\theta, s)$, so local arguments may be reduced to the fixed direction θ^*.

Proposition 3.3. ***Continuity and decay.*** *Let $f \in C(\mathbb{R}^n)$ be such that $|f(x)| = O(|x|^{-a})$ for some $a > n-1$. Then $\mathcal{R}f \in C(Z)$ and $|\mathcal{R}f(\theta, s)| = O\big(|s|^{n-1-a}\big)$. In particular, if f is rapidly decreasing on $\mathbb{R}^n$, then $\mathcal{R}f$ is rapidly decreasing on Z.*

Proof. The estimate is obtained by integrating in polar coordinates over $\theta^\perp$.

$$|\mathcal{R}f(\theta, s)| = \left| \int_{\theta^\perp} f\,(s\theta + y)\, dy \right| \overset{|f(x)| \le C(1+|x|)^{-a}}{\downarrow} \le C \int_{\theta^\perp} (1 + |s\theta + y|)^{-a}\, dy$$

$$\overset{|s\theta+y|^2 = s^2 + |y|^2}{\downarrow} = C \int_{\theta^\perp} \left(1 + \sqrt{s^2 + |y|^2}\right)^{-a} dy \underset{\substack{\uparrow \\ y = r\omega \\ \omega \in S^{n-2}_\theta}}{=} C \int_0^\infty \left(1 + \sqrt{s^2 + r^2}\right)^{-a} r^{n-2}\, dr,$$

[11] Indeed, after polar coordinates in $\theta^\perp$, the tail is bounded by $\int^\infty r^{n-2} r^{-a}\, dr < \infty$.

where $S_\theta^{n-2} := \theta^\perp \cap S^{n-1}$. For $|s| \geq 1$, the change $r = |s|u$ gives

$$|\mathcal{R}_\theta f(s)| \leq C|s|^{n-1-a} \int_0^\infty u^{n-2} \left(1+u^2\right)^{-a/2} du \overset{a>n-1}{\underset{\downarrow}{<}} \infty.^{12}$$

Hence $|\mathcal{R}_\theta f(s)| = O\big(|s|^{n-1-a}\big)$.

For the continuity, fix $(\theta_0, s_0) \in Z$. By orthogonal invariance, we may assume, after reordering coordinates if necessary, that $|\theta_n| \geq c > 0$ in a neighborhood of (θ_0, s_0). Writing $x = (x', x_n) \in \mathbb{R}^{n-1} \times \mathbb{R}$ and $\theta = (\theta', \theta_n) \in S^{n-1}$, the hyperplane θ_s is the graph

$$x_n = \varphi_{\theta,s}(x') := \frac{s - \theta' \cdot x'}{\theta_n} = \frac{s - \sum_{i=1}^{n-1} \theta_i x_i}{\theta_n},$$

Using that $|\theta'|^2 + \theta_n^2 = 1$, we get that the graph measure is $\sqrt{1 + |\nabla \varphi_{\theta,s}(x')|^2}\, dx' = \frac{dx'}{|\theta_n|}$. Therefore we may write

$$\mathcal{R}f(\theta, s) = \int_{\theta_s} f(x)\, dx = \frac{1}{|\theta_n|} \int_{\mathbb{R}^{n-1}} f\left(x', \frac{s - \theta' \cdot x'}{\theta_n}\right) dx'.$$

For fixed x', the integrand is continuous in (θ, s). Moreover, using $|\theta_n| \geq c$, $\big|(x', \varphi_{\theta,s}(x'))\big| \geq |x'|$, and our hypothesis on f, we get the uniform bound

$$\frac{1}{|\theta_n|} \left| f\left(x', \frac{s - \theta' \cdot x'}{\theta_n}\right) \right| \leq \frac{C}{c} (1 + |x'|)^{-a}. \tag{9}$$

The right-hand side of (9) is integrable on $\mathbb{R}^{n-1}$, since $a > n-1$. Therefore the theorem on continuity of parameter-dependent integrals gives continuity near (θ_0, s_0), and the arbitrariness of this point gives $\mathcal{R}f \in C(Z)$. The rapidly decreasing case follows by taking a arbitrarily large. □

Proposition 3.4. ***Smoothness.*** *Suppose that $f \in C^1(\mathbb{R}^n)$ satisfies $|f(x)| = O(|x|^{-a})$ for some $a > n-1$ and $|D_j f(x)| = O(|x|^{-b})$, $j = 1, \ldots, n$, for $b > n$. Let $\mathscr{R}f$ be the homogeneous extension of degree -1 of $\mathcal{R}f$, defined by*

$$\mathscr{R}f(\theta, s) := |\theta|^{-1} \mathcal{R}f\left(e_\theta, t\right), \quad \text{where } \theta \in \mathbb{R}^n \setminus \{0\},\ e_\theta := \frac{\theta}{|\theta|} \text{ and } t := \frac{s}{|\theta|}.$$

Then $\mathscr{R}f \in C^1((\mathbb{R}^n \setminus \{0\}) \times \mathbb{R})$, and

$$\partial_s \mathscr{R}f(\theta, s) = |\theta|^{-3} \sum_{j=1}^n \theta_j \mathcal{R}(D_j f)\left(e_\theta, t\right). \tag{10}$$

Moreover,

$$\partial_{\theta_j} \mathscr{R}f(\theta, s) = -\partial_s \mathscr{R}(x_j f)(\theta, s), \qquad j = 1, \ldots, n. \tag{11}$$

[12] Near 0 the integrand behaves like u^{n-2}, which is integrable. For $u \geq 1$, $u^{n-2}(1+u^2)^{-a/2} \leq u^{n-2}(u^2+u^2)^{-a/2} = u^{n-2}u^{-a} = u^{n-2-a}$. Since $a > n-1$, we have $n-2-a < -1$, hence $\int_1^\infty u^{n-2-a}\, du < \infty$.

Proof. Fix $e_\theta \in S^{n-1}$. Since $b > n$, each derivative $D_k f$, $k = 1, \dots, n$ satisfies the decay assumption of Prop. 3.3. Hence $\mathcal{R}(D_k f) \in C(Z)$.

Now, for $r \in \mathbb{R}$, we have $\mathcal{R}f(e_\theta, r) = \int_{e_\theta^\perp} f(re_\theta + y)\, dy$. For each $y \in e_\theta^\perp$, the map $r \mapsto f(re_\theta + y)$ is C^1 and

$$\left| \frac{d}{dr} f(re_\theta + y) \right| = \left| \sum_{j=1}^{n} (e_\theta)_j D_j f(re_\theta + y) \right| \le C\,(1 + |re_\theta + y|)^{-b} \in L^1(e_\theta^\perp).^{13}$$

Thus, we can use differentiation under the integral sign, which gives

$$\frac{\partial}{\partial r} \mathcal{R}f(e_\theta, r) = \int_{e_\theta^\perp} \frac{d}{dr} f(re + y)\, dy = \int_{e_\theta^\perp} \sum_{k=1}^{n} e_k\, D_k f(re + y)\, dy = \sum_{j=1}^{n} (e_\theta)_j \mathcal{R}(D_j f)(e_\theta, r).$$

Applying the chain rule to $\mathscr{R}f(\theta, s)$, and not forgetting that $\frac{dt}{ds} = \frac{1}{|\theta|}$, we get (10).

For the derivatives with respect to θ_j, we use the same local parametrization as in the proof of Prop. 3.3, but now We will take, first, the branch $\theta_n > 0$ (the case $\theta_n < 0$ is similar) and rewrite

$$\mathscr{R}f(\theta, s) = \frac{1}{\theta_n} \int_{\mathbb{R}^{n-1}} f\left(x', \frac{s - \theta' \cdot x'}{\theta_n}\right) dx'.$$

The derivatives of the integrand are:

$$\frac{\partial}{\partial \theta_j} \left[\frac{1}{\theta_n} f\left(x', \frac{s - \theta' \cdot x'}{\theta_n}\right) \right] = -\frac{x_j}{\theta_n^2}\, D_n f\left(x', \frac{s - \theta' \cdot x'}{\theta_n}\right), \quad j = 1, \dots, n-1;$$

$$\frac{\partial}{\partial \theta_n} \left[\frac{1}{\theta_n} f\left(x', \frac{s - \theta' \cdot x'}{\theta_n}\right) \right] = -\frac{1}{\theta_n^2} f\left(x', \frac{s - \theta' \cdot x'}{\theta_n}\right) - \frac{s - \theta' \cdot x'}{\theta_n^3} D_n f\left(x', \frac{s - \theta' \cdot x'}{\theta_n}\right).$$

The first term is controlled by the same estimate used in the proof of Prop. 3.3. The new terms are those involving $D_n f$. In the present local chart, s, θ and $1/\theta_n$ are bounded, and therefore

$$|x_j| \le |x'|, \quad j < n, \qquad \left| \frac{s - \theta' \cdot x'}{\theta_n} \right| \le C(1 + |x'|).$$

Hence the terms involving $D_n f$ are bounded by a constant multiple of $(1 + |x'|)(1 + |x|)^{-b} \le (1 + |x'|)^{1-b}$, which is integrable on $\mathbb{R}^{n-1}$, since $b > n$. Therefore the same dominated-convergence argument as before justifies differentiating under the integral sign for $j = 1, \dots, n$ and gives

$$\frac{\partial \mathscr{R}f}{\partial \theta_j}(\theta, s) = -\frac{1}{\theta_n^2} \int_{\mathbb{R}^{n-1}} x_j\, D_n f\left(x', \frac{s - \theta' \cdot x'}{\theta_n}\right) dx' \quad j = 1, \dots, n-1;$$

$$\frac{\partial \mathscr{R}f}{\partial \theta_n}(\theta, s) = -\frac{1}{\theta_n^2} \int_{\mathbb{R}^{n-1}} f\left(x', \frac{s - \theta' \cdot x'}{\theta_n}\right) dx' - \frac{1}{\theta_n^2} \int_{\mathbb{R}^{n-1}} x_n\, D_n f\left(x', \frac{s - \theta' \cdot x'}{\theta_n}\right) dx',$$

[13] Since $y \in e_\theta^\perp$, $|re_\theta + y|^2 = r^2 + |y|^2 \ge |y|^2$. Hence $C(1 + |re_\theta + y|)^{-b} \le C(1 + |y|)^{-b}$ and $\int_{e_\theta^\perp} (1 + |y|)^{-b}\, dy < \infty$ because $e_\theta^\perp \simeq \mathbb{R}^{n-1}$ and $b > n - 1$.

where we used, for $\partial\theta_n$, that $x_n = \frac{s-\theta'\cdot x'}{\theta_n}$, from the parametrization.

We are left with $\partial_s \mathscr{R}(x_j f)(\theta, s)$. To deal with it, we have $D_k x_j f(x) = \delta_{jk} f(x) + x_j D_k f(x)$. For $j < n$, we have $|D_k x_j f(x)| = O(|x|^{-a}) + O(|x|^{1-b})$ and thus the first part of the proof applies to $x_j f$, and $\dfrac{\partial}{\partial s}\mathscr{R}(x_j f)$ exists and is continuous. Since the coordinate x_j does not depend on s this parametrization, we have

$$\mathscr{R}(x_j f)(\theta, s) = \frac{1}{\theta_n}\int_{\mathbb{R}^{n-1}} x_j\, f\left(x', \frac{s-\theta'\cdot x'}{\theta_n}\right) dx',$$
$$\frac{\partial \mathscr{R}(x_j f)}{\partial s}(\theta, s) = \frac{1}{\theta_n^2}\int_{\mathbb{R}^{n-1}} x_j\, D_n f\left(x', \frac{s-\theta'\cdot x'}{\theta_n}\right) dx'.$$

We have proved (11) for $j < n$. For $\partial_s \mathscr{R}(x_n f)(\theta, s)$ the procedure is similar but taking into account that $\frac{\partial x_n}{\partial s} = \frac{1}{\theta_n}$. □

Corollary 3.5. ***Preservation of*** $\mathcal{S}$. *The Radon transform maps* $\mathcal{S}(\mathbb{R}^n)$ *into* $\mathcal{S}(Z)$.

Most of the results on the Radon transform discussed for the rest of the work hold for continuous or smooth functions with suitable decay at infinity [CMOCP07, Secs. 2.1–2.4]. However, in order to simplify the exposition, from now on we shall mainly work with functions in the Schwartz space $\mathcal{S}(\mathbb{R}^n)$.

Theorem 3.6. ***(Projection–slice).*** *If* $f \in \mathcal{S}(\mathbb{R}^n)$, *then* $\widehat{\mathcal{R}f}(\theta, \sigma) = \widehat{f}(\sigma\theta)$, *for* $\theta \in S^{n-1}$, $\sigma \in \mathbb{R}$.

Proof.
$$\begin{aligned}\widehat{\mathcal{R}f}(\theta,\sigma) &= \int_{\mathbb{R}} e^{-2\pi i s\sigma}\mathcal{R}f(\theta, s)\, ds = \int_{\mathbb{R}} e^{-2\pi i s\sigma}\left(\int_{\theta^\perp} f(s\theta + y)\, dy\right) ds \\ &= \int_{\mathbb{R}}\int_{\theta^\perp} e^{-2\pi i s\sigma} f(s\theta + y)\, dy\, ds \\ &= \int_{\mathbb{R}^n} e^{-2\pi i(\sigma\theta)\cdot x} f(x)\, dx = \widehat{f}(\sigma\theta). \quad \square\end{aligned}$$

Corollary 3.7. *The Radon transform* $\mathcal{R} : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}(Z)$ *is injective.*

Proof. If $\mathcal{R}f = 0$, then $\widehat{\mathcal{R}f}(\theta, \sigma) = 0$. By Theorem 3.6, $\widehat{f}(\sigma\theta) = 0$ for all $\theta \in S^{n-1}$, $\sigma \in \mathbb{R}$. Every $\xi \in \mathbb{R}^n$ can be written as $\xi = \sigma\theta$. Hence $\widehat{f} \equiv 0$, and the injectivity of the FT gives $f = 0$. □

Corollary 3.8. *If* $f_1, f_2 \in \mathcal{S}(\mathbb{R}^n)$, *then* $\mathcal{R}(f_1 * f_2) = (\mathcal{R}f_1) *_s (\mathcal{R}f_2)$.

Proof.
$$\widehat{\mathcal{R}(f_1 * f_2)}(\theta,\sigma) \overset{\text{Thm. 3.6}}{=} \widehat{f_1 * f_2}(\sigma\theta) = \widehat{f_1}(\sigma\theta)\,\widehat{f_2}(\sigma\theta) = \widehat{\mathcal{R}f_1}(\theta,\sigma)\,\widehat{\mathcal{R}f_2}(\theta,\sigma)$$
$$= \widehat{\mathcal{R}f_1 *_s \mathcal{R}f_2}(\theta,\sigma).$$
Applying $\mathcal{F}_s^{-1}$ gives $\mathcal{R}(f_1 * f_2)(\theta, s) = (\mathcal{R}f_1 * \mathcal{R}f_2)(\theta, s)$. □

3.2 Uniqueness

We have just proved a global uniqueness result for $\mathcal{R}f$: by Corollary 3.7, $\mathcal{R}f(Z)$ determines the function, at least in the class $\mathcal{S}(\mathbb{R}^n)$. In other words, if $\mathcal{R}f$ is known on the whole space of hyperplanes Z then f is uniquely determined.

However, in many applications $\mathcal{R}f$ is not known on all of Z. One may know it only for some directions, or only for hyperplanes intersecting a certain region. We ask whether these partial data determine the function, or at least determine its restriction to a certain subset of $\mathbb{R}^n$. This leads to the question of local uniqueness. The next two results address this type of question.

Theorem 3.9. ***Support-hole.*** *Let $f \in \mathcal{S}(\mathbb{R}^n)$, and let $K \subset \mathbb{R}^n$ be a compact convex set. Assume that $\mathcal{R}f(\theta, s) = 0$ for every hyperplane θ_s which does not meet K. Then $f(x) = 0$ for $x \in \mathbb{R}^n \setminus K$.*

We only consider the compactly supported case, that is, $f \in C_c^\infty(\mathbb{R}^n)$. In this case, we are going to follow Strichartz's approach [Str82, Sec. 2], which considerably simplifies the general proof due to Helgason [CMOCP07, Sec. 4.3]. Also to make the geometric idea easier to visualize, in the proof below we only prove the case, $f \in C_c^\infty(\mathbb{R}^2)$. The same idea can be carried out in arbitrary dimension.[14]

Proof. We prove the case $n = 2$ and $f \in C_c^\infty(\mathbb{R}^2)$. Let $x_0 \in \mathbb{R}^2 \setminus K$. By the separating hyperplane theorem[15], there are $\theta \in S^1$ and $\tilde{s} \in \mathbb{R}$ such that $x_0 \in \theta_{\tilde{s}}$, and $\theta_{\tilde{s}} \cap K = \emptyset$. Choosing the orientation of θ, we may assume that the parallel lines θ_s, with $s \geq \tilde{s}$, lie on the right side of K. Hence, $\theta_s \cap K = \emptyset$ if $s \geq \tilde{s}$. After a rotation of coordinates, we may assume that $\theta = \theta^* := (0, 1)$. Thus $x_0 \in \theta^*_{\tilde{s}}$, with $\theta^*_s \cap K = \emptyset$ for $s \geq \tilde{s}$.

We first explain why it is enough to prove that polynomially weighted Radon data also vanish. Concretely, suppose that

$$\mathcal{R}(Pf)(\omega, \sigma) = 0 \quad \text{for any polynomial } P, \text{ for every line } \omega_\sigma \subset K^c. \tag{12}$$

Assume this is true for our particular line $\theta^*_{\tilde{s}} \subset K^c$, which contains x_0. Let $E := \theta^*_{\tilde{s}} \cap \operatorname{supp}(f)$. This set is compact. By the Weierstrass approximation theorem, there are polynomials P_m such that $P_{m|E} \longrightarrow \overline{f}_{|E}$ uniformly on E. Then

$$0 \overset{(12)}{=} \mathcal{R}(P_m f)(\theta^*, \tilde{s}) = \int_{\theta^*_{\tilde{s}}} P_m(x) f(x)\, dx = \int_E P_m(x) f(x)\, dx \longrightarrow \int_E |f(x)|^2\, dx.$$

[14] In the n-dimensional compactly supported case one first uses the fact that a compact convex set in $\mathbb{R}^n$ is the intersection of the closed balls which contain it [GJSM04, Thm. 1.1 and Prop. 1.3]. Hence it is enough to prove the result when K is a closed ball. After this reduction, the proof follows the same strategy as above: one proves the vanishing of the polynomially weighted Radon data, by rotating the hyperplane, differentiating the resulting identity and integrating in the normal variable.

[15] Recall that, in finite dimension, a point outside a closed convex set can be separated from it by an affine hyperplane.

Hence $\int_E |f(x)|^2\,dx = 0$. Since f is continuous, this implies $f = 0$ on E. In particular, because $x_0 \in \theta^*_{\tilde{s}}$, we get $f(x_0) = 0$ and we prove the theorem.

Thus, it remains to prove (12). In order to do so, we make a claim: for every linear polynomial $q(y) = a + by$, it holds

$$0 = a\int_{\mathbb{R}} f(y,\tilde{s})\,dy + b\int_{\mathbb{R}} yf(y,\tilde{s})\,dy = \int_{\mathbb{R}} q(y)f(y,\tilde{s})\,dy. \tag{13}$$

Now we prove the claim. The first term in (13) is zero by hypothesis, since for $s \geq \tilde{s}$, we have

$$\mathcal{R}f(\theta^*, s) = \int_{(\theta^*)^\perp} f\big(s\theta^* + y\big)\,dy = \int_{\mathbb{R}} f(y,s)\,dy = 0, \qquad s \geq \tilde{s}. \tag{14}$$

To show
$$\int_{\mathbb{R}} y\,f(y,\tilde{s})\,dy = 0, \tag{15}$$

We will keep an $s \geq \widetilde{s}$ fixed and rotate the normal vector of the line. This will replace θ^* by nearby vectors $\theta_\alpha = (\sin\alpha, \cos\alpha) \in S^1$ with α small enough so that the slightly rotated line $(\theta_\alpha)_s$ still does not meet K. We are not moving the line in parallel (that would mean changing s). Recall that $(\theta_\alpha)_s = \{x \in \mathbb{R}^2 : x \cdot \theta_\alpha = s\}$ by definition, so $\theta_\alpha^\perp = \{y(\cos\alpha, -\sin\alpha) : y \in \mathbb{R}\}$. Hence

$$\begin{aligned} 0 \overset{(\theta_\alpha)_s \cap K = \emptyset}{\downarrow} &\mathcal{R}f(\theta_\alpha, s) = \int_{\theta_\alpha^\perp} f(s\theta_\alpha + u)\,du = \int_{\mathbb{R}} f\big(s\theta_\alpha + y(\cos\alpha, -\sin\alpha)\big)\,dy \\ &= \int_{\mathbb{R}} f(s\sin\alpha + y\cos\alpha,\ s\cos\alpha - y\sin\alpha)\,dy. \end{aligned}$$

Differentiating this identity at $\alpha = 0$, under the integral sign, we obtain[16]

$$0 = \int_{\mathbb{R}} yD_2f(y,s)\,dy = 0, \qquad s \geq \tilde{s}. \tag{16}$$

Now to obtain (15), we integrate (16) in s:

$$0 = \int_{\tilde{s}}^{\infty}\int_{\mathbb{R}} yD_2f(y,s)\,dy\,ds \overset{\text{Fubini}}{\downarrow} \int_{\mathbb{R}} y\,[f(y,s)]_{s=\tilde{s}}^{s=+\infty}\,dy \overset{f\in C_c^\infty(\mathbb{R}^2)}{\downarrow} -\int_{\mathbb{R}} y\,f(y,\tilde{s})\,dy.$$

Putting (14) and (15) together proves our claim (13). Now using (13), we are able to prove (12). We will prove (12) using induction on $d_P := \deg P$. For $d_P = 0$,

[16]
$$\begin{aligned} 0 &= \frac{d}{d\alpha}\Big|_{\alpha=0}\int_{\mathbb{R}} f(s\sin\alpha + y\cos\alpha,\ s\cos\alpha - y\sin\alpha)\,dy \\ &= \int_{\mathbb{R}}\left[D_1f(y,s)\frac{d}{d\alpha}\Big|_{\alpha=0}(s\sin\alpha + y\cos\alpha) + D_2f(y,s)\frac{d}{d\alpha}\Big|_{\alpha=0}(s\cos\alpha - y\sin\alpha)\right]dy \\ &= \int_{\mathbb{R}} (sD_1f(y,s) - yD_2f(y,s))\,dy \overset{*}{=} \int_{\mathbb{R}} yD_2f(y,s)\,dy = 0, \qquad s \geq \tilde{s}, \end{aligned}$$

where $*$ is because the first term vanishes, since $f \in C_c^\infty(\mathbb{R}^2)$.

P is a constant and so $\mathcal{R}(Pf)(\omega,\sigma) = 0$, $\forall \omega_\sigma \subset K^c$. For $d_p \leq m$, put $g := pf \in C_c^\infty(\mathbb{R}^2)$, and assume that $\mathcal{R}(g)(\omega,\sigma) = 0$ for every $\omega_\sigma \subset K^c$. Take P with $d_P = m+1$. By linearity, it is enough to consider a monomial, and we may write $P = q\,p$, where $d_q = 1$ and $d_p \leq m$. This gives $\mathcal{R}(Pf)(\omega,\sigma) = \mathcal{R}(qpf)(\omega,\sigma) = \mathcal{R}(qg)(\omega,\sigma) = 0$ for every line ω_σ which does not meet K. This proves (12). □

The next result deals with the opposite type of partial data: instead of knowing that $\mathcal{R}f$ vanishes outside a region for every direction, we assume that it vanishes for every $s \in \mathbb{R}$, but only for a prescribed set of directions. The relevant condition is that this set of directions determines homogeneous polynomials.

Definition 3.10. *A function $p_k : \mathbb{R}^n \to \mathbb{C}$ is called a **homogeneous polynomial of degree** k if it can be written*

$$p_k(x) = \sum_{|\alpha|=k} c_\alpha x^\alpha, \qquad x \in \mathbb{R}^n,$$

where $c_\alpha \in \mathbb{C}$, $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a multi-index and $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. We denote by $\mathcal{P}_k^{\hom}(\mathbb{R}^n)$ the complex vector space of such polynomials, together with the zero polynomial.

If $p_k \in \mathcal{P}_k^{\hom}(\mathbb{R}^n)$ then $p_k(\lambda x) = \lambda^k p_k(x)$, for $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$.[17] Notice that $p_k \in \mathcal{P}_k^{\hom}(\mathbb{R}^n)$ implies that p_k is even if k is even and it is odd if k is odd. In particular, for $\theta \in S^{n-1}$, $p_k(\sigma\theta) = \sigma^k p_k(\theta)$, for $\sigma \in \mathbb{R}$.

Definition 3.11. *Let $\Theta \subset S^{n-1}$. We say that Θ is a **uniqueness set for homogeneous polynomials** in $\mathbb{R}^n$ if, for every homogeneous polynomial $p : \mathbb{R}^n \to \mathbb{C}$, we have that $p(\theta) = 0$ for all $\theta \in \Theta$ implies $p \equiv 0$ in $\mathbb{R}^n$.*

Theorem 3.12. ***Uniqueness from a set of directions.*** *Let $\Theta \subset S^{n-1}$ be a uniqueness set for homogeneous polynomials. If $f \in C_c(\mathbb{R}^n)$ satisfies $\mathcal{R}f(\theta,s) = 0$ for every $\theta \in \Theta$ and every $s \in \mathbb{R}$, then $f \equiv 0$.*

Proof. Since $f \in C_c(\mathbb{R}^n)$, $\widehat{f}$ extends to an entire function F on $\mathbb{C}^n$ [18]. In particular, $\widehat{f} = F_{|\mathbb{R}^n}$ is real analytic[19], and near the origin it admits the form of a multivariable

[17] Indeed, if $p_k(x) = \sum_{|\alpha|=k} c_\alpha x^\alpha$ and $|\alpha| = \alpha_1 + \cdots + \alpha_n = k$, then

$$p_k(\lambda x) = \sum_{|\alpha|=k} c_\alpha (\lambda x)^\alpha = \sum_{|\alpha|=k} c_\alpha (\lambda x_1)^{\alpha_1} \cdots (\lambda x_n)^{\alpha_n} = \sum_{|\alpha|=k} c_\alpha \lambda^{\alpha_1+\cdots+\alpha_n} x^\alpha = \sum_{|\alpha|=k} c_\alpha \lambda^k x^\alpha = \lambda^k p_k(x).$$

[18] Namely, $F(z) = \int_{\operatorname{supp} f} f(x) e^{-2\pi i x\cdot z}\, dx$. The proof is the same dominated-convergence argument as for the Paley–Wiener representative.

[19] Real analytic means that locally the function coincides pointwise with an absolutely convergent power series in the real variables. Since holomorphic functions are locally given by convergent power series, the restriction of an entire function on $\mathbb{C}^n$ to $\mathbb{R}^n$ is real analytic.

Taylor expansion $\widehat{f}(\zeta) = \sum_{k=0}^{\infty} p_k(\zeta)$[20] where each p_k is a homogeneous polynomial of degree k.

Fix $\theta \in \Theta$. By hypothesis, $\mathcal{R}f(\theta, s) = 0$ for $s \in \mathbb{R}$. Using Thm. 3.6, we obtain $0 = \widehat{\mathcal{R}f}(\theta, \sigma) = \widehat{f}(\sigma\theta)$ for $\sigma \in \mathbb{R}$. Substituting the homogeneous expansion of $\widehat{f}$

$$0 = \widehat{f}(\sigma\theta) = \sum_{k=0}^{\infty} p_k(\sigma\theta) = \sum_{k=0}^{\infty} \sigma^k p_k(\theta), \qquad \sigma \in \mathbb{R}.$$

Thus, for each fixed $\theta \in \Theta$, all coefficients of this power series in σ vanish. Hence $p_k(\theta) = 0$ for every $\theta \in \Theta$, $k \geq 0$. Since Θ is a uniqueness set for homogeneous polynomials, it follows that $p_k \equiv 0$ for every k. Therefore all Taylor coefficients of $\widehat{f}$ at the origin vanish. Since $\widehat{f}$ is real analytic, it coincides with its Taylor series near 0; hence $\widehat{f} = 0$ in an open neighborhood of 0. By analytic continuation, $\widehat{f} \equiv 0$ on $\mathbb{R}^n$. Finally, by the injectivity of the FT, $f \equiv 0$. □

The two uniqueness results use complementary Radon data. In Theorem 3.9, the vanishing is assumed for every direction $\theta \in S^{n-1}$, but only for those values of s such that the hyperplane θ_s avoids K. By contrast, in Theorem 3.12, the vanishing is assumed only for directions in a prescribed set $\Theta \subset S^{n-1}$, but for every $s \in \mathbb{R}$. If Θ is rich enough to determine homogeneous polynomials, then this restricted directional information is already enough to force $f \equiv 0$.

3.3 Range

We now describe which functions in $\mathcal{S}(Z)$ arise as Radon transforms of functions in $\mathcal{S}(\mathbb{R}^n)$. Since we want to reconstruct f from $\mathcal{R}f$, we have to know if $\mathcal{R}(\mathcal{S}(\mathbb{R}^n))$ is the whole $\mathcal{S}(Z)$, or a smaller class. We will see that the map $\mathcal{R}$ is not onto: by Prop. 3.2, the Radon transform is even $\mathcal{R}f(-\theta, -s) = \mathcal{R}f(\theta, s)$; second, there is a restriction on the moments in the s-variable.

Notation: if $g \in \mathcal{S}(Z)$, its m-th moment is $\mu_m(\theta) := \int_{\mathbb{R}} s^m g(\theta, s)\, ds$ for $\theta \in S^{n-1}$.

If g is actually of the form $g = \mathcal{R}f$, then each of its moments $\mu_m(\theta)$ must be the restriction to S^{n-1} of some $p_m \in \mathcal{P}_m^{\hom}(\mathbb{R}^n)$. This restriction is important enough to have its own name.

Definition 3.13. *Let $g \in \mathcal{S}(Z)$. We say that g satisfies the **Helgason–Ludwig consistency conditions**, abbreviated* HLCC, *if, for every $m \in \mathbb{N} \cup \{0\}$, there exists a polynomial $p_m \in \mathcal{P}_m^{\hom}(\mathbb{R}^n)$ such that $\mu_m(\theta) = p_m(\theta)$ for $\theta \in S^{n-1}$.*

[20] $\widehat{f}(\zeta) = \sum_{\alpha \in \mathbb{N}_0^n} \frac{1}{\alpha!}\, \partial^\alpha \widehat{f}(0)\, \zeta^\alpha = \sum_{\substack{\alpha \in \mathbb{N}_0^n \\ |\alpha|=0}} \frac{1}{\alpha!}\, \partial^\alpha \widehat{f}(0)\, \zeta^\alpha + \sum_{\substack{\alpha \in \mathbb{N}_0^n \\ |\alpha|=1}} \frac{1}{\alpha!}\, \partial^\alpha \widehat{f}(0)\, \zeta^\alpha + \cdots$

$$= \sum_{k=0}^{\infty} \left(\sum_{\alpha \in \mathbb{N}_0^n,\, |\alpha|=k} \frac{1}{\alpha!}\, \partial^\alpha \widehat{f}(0)\, \zeta^\alpha \right) = \sum_{k=0}^{\infty} p_k(\zeta),$$

These two restrictions are not merely necessary on $\mathcal{R}(\mathcal{S}(\mathbb{R}^n))$. They actually characterize it. As a bonus, an analogous statement holds for $C_c^\infty(\mathbb{R}^n)$.

Theorem 3.14. ***Exact range of the Radon transform.*** *The range of $\mathcal{R} : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}(Z)$ is exactly the space of all even functions $g \in \mathcal{S}(Z)$ satisfying the* HLCC.

Proof. Let $f \in \mathcal{S}(\mathbb{R}^n)$ and set $g = \mathcal{R}f$. We already know from 3.5 and 3.2 that $g \in \mathcal{S}(Z)$ and that g is even. We check the moment conditions. For $m \in \mathbb{N} \cup \{0\}$ and $\theta \in S^{n-1}$,

$$\begin{aligned}\int_{\mathbb{R}} s^m g(\theta, s)\, ds &= \int_{\mathbb{R}} \int_{\theta^\perp} s^m f(s\theta + y)\, dy\, ds \overset{\substack{y\in\theta^\perp\\ \downarrow}}{=} \int_{\mathbb{R}} \int_{\theta^\perp} \big((s\theta + y)\cdot\theta\big)^m f(s\theta + y)\, dy\, ds \\ &\overset{\substack{x=s\theta+y\\ \downarrow}}{=} \int_{\mathbb{R}^n} (x\cdot\theta)^m f(x)\, dx \overset{*}{=} \int_{\mathbb{R}^n} \left(\sum_{|\alpha|=m} \tfrac{m!}{\alpha!}\, x^\alpha \theta^\alpha \right) f(x)\, dx \\ &= \sum_{|\alpha|=m} \tfrac{m!}{\alpha!} \left(\int_{\mathbb{R}^n} x^\alpha f(x)\, dx \right) \theta^\alpha = p_m(\theta),\end{aligned}$$

where we have used the multinomial formula[21] in $*$ and

$$p_m(y) := \sum_{|\alpha|=m} \tfrac{m!}{\alpha!} \left(\int_{\mathbb{R}^n} x^\alpha f(x)\, dx \right) y^\alpha, \qquad y \in \mathbb{R}^n.$$

The integrals are finite because f is a Schwartz function, and p_m is homogeneous of degree m, unless it is identically zero. This proves the necessary conditions.

Conversely, let $g \in \mathcal{S}(Z)$ be even and satisfy the *HLCC*. If $\mathcal{R}f = g$, then Thm. 3.6 gives for $\sigma > 0$, $\widehat{f}(\sigma\theta) = \widehat{g}(\theta, \sigma)$. Thus, the natural candidate for $\widehat{f}$ is $F(\sigma\theta) := \widehat{g}(\theta, \sigma)$.

It is enough to prove that F extends to a function in $\mathcal{S}(\mathbb{R}^n)$: since the FT is a homeomorphism of $\mathcal{S}(\mathbb{R}^n)$ onto itself, such an extension would be the FT of some $f \in \mathcal{S}(\mathbb{R}^n)$.

We write ξ for the Cartesian variable in $\mathbb{R}^n$. Since $\widehat{g} \in \mathcal{S}(Z)$, the function

$$F(\xi) = F(\sigma\theta) = \widehat{g}(\theta, \sigma), \qquad \xi = \sigma\theta, \quad \sigma > 0,$$

is smooth on $\mathbb{R}^n \setminus \{0\}$, and the Schwartz estimates away from the origin follow from the rapid decay of $\widehat{g}$ in σ. Hence the only point to control is $\xi = 0$.

[21] For $x, y \in \mathbb{R}^n$, $m \in \mathbb{N}$, and multiindices $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$,

$$(x\cdot y)^m = \left(\sum_{j=1}^n x_j y_j \right)^m = \sum_{|\alpha|=m} \frac{m!}{\alpha!} x^\alpha y^\alpha, \qquad \alpha! := \alpha_1! \cdots \alpha_n! \quad x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}.$$

We prove that all Cartesian derivatives of F are locally bounded near the origin. Then, by an extension argument, this will be enough to conclude that $F \in \mathcal{S}(\mathbb{R}^n)$. We first write $\partial/\partial\xi_j$ in polar coordinates.

Let $\xi = \sigma\theta$, with $\sigma = |\xi| > 0$ and $\theta = \xi/|\xi|$. Then, for $1 \le j, \ell \le n$, by the chain rule,

$$\frac{\partial}{\partial \xi_j} = \theta_j \frac{\partial}{\partial \sigma} + \frac{1}{\sigma} \sum_{\ell=1}^{n} (\delta_{j\ell} - \theta_j \theta_\ell) \frac{\partial}{\partial \theta_\ell}.$$[22]

Applying this identity repeatedly, each differentiation with respect to ξ produces derivatives with respect to σ and to the angular variables, and introduces at most one additional factor $1/\sigma$. Thus, by induction on $q = |\alpha|$, one obtains

$$\frac{\partial^q}{\partial \xi^\alpha} = \sum_{1 \le |\beta| + k \le q} \frac{c_{\beta,k}(\theta)}{\sigma^{q-k}} \frac{\partial^{|\beta|+k}}{\partial \theta^\beta \partial \sigma^k}, \qquad (\alpha \in \mathbb{N}^n,\ |\alpha| = q \ge 1),$$

where the $c_{\beta,k}$ are polynomials in θ, hence uniformly bounded on S^{n-1}. Hence, singularities at the origin can only come from the powers of σ^{-1}.

To see that these singularities cancel, write the Taylor decomposition

$$e(u) := e^{-2\pi i u} = \sum_{m=0}^{q-1} \tfrac{(-2\pi i u)^m}{m!} + e_q(u). \tag{17}$$

By the *HLCC* and the homogeneity of p_m, we have

$$\sigma^m \int_{\mathbb{R}} s^m g(\theta, s)\, ds \overset{\text{HLCC}}{\underset{\downarrow}{=}} \sigma^m p_m(\theta) \overset{\text{homogeneity}}{\underset{\downarrow}{=}} p_m(\sigma\theta), \tag{18}$$

therefore

$$\begin{aligned} F(\sigma\theta) = \int_{\mathbb{R}} e^{-2\pi i \sigma s} g(\theta, s)\, ds &\overset{\text{Eq. (17)}}{\underset{\text{Eq. (18)}}{=}} \sum_{m=0}^{q-1} \tfrac{(-2\pi i)^m}{m!} \sigma^m \int_{\mathbb{R}} s^m g(\theta, s)\, ds + \int_{\mathbb{R}} e_q(\sigma s) g(\theta, s)\, ds \\ &= \sum_{m=0}^{q-1} \tfrac{(-2\pi i)^m}{m!}\, p_m(\sigma\theta) + \int_{\mathbb{R}} e_q(\sigma s)\, g(\theta, s)\, ds = P_q(\sigma\theta) + \int_{\mathbb{R}} e_q(\sigma s) g(\theta, s)\, ds, \end{aligned}$$

where P_q is a polynomial of degree $< q$. Hence $\partial_\xi^\alpha P_q = 0$ whenever $|\alpha| = q$. Moreover, e_q does not depend on θ, while g does not depend on σ. So, differentiating, we are left with

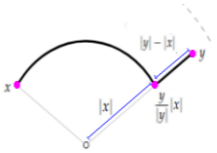
$$\begin{aligned} \partial_\xi^\alpha F(\sigma\theta) = \partial_\xi^\alpha \left(\int_{\mathbb{R}} e_q(\sigma s)\, g(\theta, s)\, ds \right) &= \sum_{1 \le |\beta| + k \le q} \tfrac{c_{\beta,k}(\theta)}{\sigma^{q-k}} \partial_\theta^\beta \partial_\sigma^k \left(\int_{\mathbb{R}} e_q(\sigma s)\, g(\theta, s)\, ds \right) \\ &\overset{*}{=} \sum_{1 \le |\beta| + k \le q} \tfrac{c_{\beta,k}(\theta)}{\sigma^{q-k}} \int_{\mathbb{R}} \partial_\theta^\beta \partial_\sigma^k \big(e_q(\sigma s)\, g(\theta, s) \big)\, ds \\ &= \sum_{1 \le |\beta| + k \le q} \int_{\mathbb{R}} \tfrac{c_{\beta,k}(\theta)}{\sigma^{q-k}} s^k e_q^{(k)}(\sigma s) \partial_\theta^\beta g(\theta, s)\, ds, \end{aligned} \tag{19}$$

[22] Using that $\dfrac{\partial \sigma}{\partial \xi_j} = \dfrac{\xi_j}{|\xi|} = \theta_j$, and that $\dfrac{\partial \theta_\ell}{\partial \xi_j} = \dfrac{\partial}{\partial \xi_j}\left(\dfrac{\xi_\ell}{|\xi|}\right) = \dfrac{\delta_{j\ell}}{\sigma} - \dfrac{\xi_\ell}{\sigma^2}\dfrac{\partial \sigma}{\partial \xi_j} = \dfrac{\delta_{j\ell} - \theta_\ell \theta_j}{\sigma}$.

where we have used differentiation under the integral sign in $*$. Using $|e_q^{(k)}(u)| \lesssim |u|^{q-k}$ for $0 \le k \le q$ [23], we can bound (19):

$$\begin{aligned}\left|\partial_\xi^\alpha F(\sigma\theta)\right| &\lesssim \sum_{1\le|\beta|+k\le q} \int_{\mathbb{R}} \tfrac{|c_{\beta,k}(\theta)|}{\sigma^{q-k}} |s|^k |\sigma s|^{q-k} \left|\partial_\theta^\beta g(\theta,s)\right| ds \\ &\overset{\sigma\ge 0}{=} \sum_{1\le|\beta|+k\le q} |c_{\beta,k}(\theta)| \int_{\mathbb{R}} |s|^q \left|\partial_\theta^\beta g(\theta,s)\right| ds \\ &\underset{|c_{\beta,k}(\theta)|\lesssim 1}{\overset{\uparrow}{\lesssim}} \sum_{1\le|\beta|+k\le q} \int_{\mathbb{R}} |s|^q \left|\partial_\theta^\beta g(\theta,s)\right| ds \underset{g\in S(Z)}{\overset{\uparrow}{\lesssim}} 1.\end{aligned}$$

Hence all Cartesian derivatives of F are locally bounded near the origin. We now explain why this implies that these derivatives have limits at the origin.



Fix a multi-index α, and choose $M = M(r) > 0$ such that $|\nabla(\partial_\xi^\alpha F)(\xi)| \le M$ when $0 < |\xi| < r$. Let $x, y \in B_r(0) \setminus \{0\}$. We join x and y by a curve $\gamma \subset B_r(0) \setminus \{0\}$ made of two pieces: first, a spherical arc joining x and $\frac{y}{|y|}|x|$, which has length smaller than $\pi|x|$, and then a radial segment joining this point with y, which has length $||x| - |y||$. This curve satisfies

$$\operatorname{long}(\gamma) \le \pi|x| + \big||x| - |y|\big| \le \pi|x| + |x| + |y| \le (\pi+1)(|x| + |y|).$$

Applying the fundamental theorem of calculus along γ, we obtain

$$\left|\partial_\xi^\alpha F(x) - \partial_\xi^\alpha F(y)\right| \le \int_\gamma \left|\nabla(\partial_\xi^\alpha F)(\xi)\right| d\ell(\xi) \le M \operatorname{long}(\gamma) \le M(\pi+1)(|x| + |y|).$$

In particular, $\left|\partial_\xi^\alpha F(x_m) - \partial_\xi^\alpha F(y_m)\right| \le MC(|x_m| + |y_m|) \xrightarrow{x_m, y_m \searrow 0} 0$. All sequences approaching 0 give the same limit, so we can define $\partial_\xi^\alpha F(0) := \lim_{\xi\to 0} \partial_\xi^\alpha F(\xi)$.

The extensions are compatible with differentiation by the same radial fundamental-theorem argument: if G_α denotes the continuous extension of $\partial_\xi^\alpha F$, then

$$G_\alpha(x) - G_\alpha(0) = \int_0^1 \sum_{j=1}^n x_j\, G_{\alpha+e_j}(tx)\, dt = \sum_{j=1}^n x_j\, G_{\alpha+e_j}(0) + o(|x|).$$

Hence $\partial_{\xi_j} G_\alpha(0) = G_{\alpha+e_j}(0)$, and the extensions are compatible with differentiation. Therefore $F \in C^\infty(\mathbb{R}^n)$.

[23] Since $e_q(u) = e^{-2\pi i u} - \sum_{m=0}^{q-1} (-2\pi i u)^m / m!$, Taylor's formula with integral remainder gives, for $0 \le k < q$,

$$e_q^{(k)}(u) = \frac{(-2\pi i)^q}{(q-k-1)!} \int_0^u \frac{(u-t)^{q-k-1}\, dt}{e^{2\pi i t}},$$

which implies $|e_q^{(k)}(u)| \le C \left|\int_0^u (u-t)^{q-k-1}\, dt\right| \le C \int_0^{|u|} r^{q-k-1}\, dr = C|u|^{q-k}$.

It remains only to check rapid decay. For $\sigma > 1$, the same polar formula gives, for every $N \in \mathbb{N}$,

$$\sigma^N |\partial_\xi^\alpha F(\sigma\theta)| \leq C \sum_{1 \leq |\beta|+k \leq q} \sigma^N \left| \partial_\theta^\beta \partial_\sigma^k \widehat{g}(\theta, \sigma) \right|.$$

The right-hand side is uniformly bounded because $\widehat{g} \in \mathcal{S}(Z)$. Hence $F \in \mathcal{S}(\mathbb{R}^n)$.

Since the FT is a homeomorphism of $\mathcal{S}(\mathbb{R}^n)$ onto itself, there exists $f \in \mathcal{S}(\mathbb{R}^n)$ such that $\widehat{f} = F$. By construction, $\widehat{f}(\sigma\theta) = \widehat{g}(\theta, \sigma)$, $\sigma > 0$. The evenness of g gives the same identity for every $\sigma \in \mathbb{R}$. Applying again Thm. 3.6, $\widehat{\mathcal{R}f}(\theta, \sigma) = \widehat{f}(\sigma\theta) = \widehat{g}(\theta, \sigma)$. By uniqueness of $\mathcal{F}_s$, $\mathcal{R}f = g$. □

Combining the range with the support theorems gives the corresponding compactly supported version. The direct implication is immediate. If $g \in C_c^\infty(Z)$ is even and satisfies the *HLCC*, Thm. 3.14 gives $f \in \mathcal{S}(\mathbb{R}^n)$ such that $\mathcal{R}f = g$. Since g vanishes for $|s|$ large, Thm. 3.9 implies that $\operatorname{supp}(f)$ is compact.

Corollary 3.15. ***Compactly supported range.*** *The range of $\mathcal{R} : C_c^\infty(\mathbb{R}^n) \longrightarrow C_c^\infty(Z)$ is the space of all even functions $g \in C_c^\infty(Z)$ satisfying* HLCC.

3.4 Inversion formulas

We now derive two inversion formulas for the $\mathcal{R}f$, i.e. formulas to obtain f from $\mathcal{R}f$. The first one expresses f in terms of $\mathcal{F}_s(\mathcal{R}f)$. The second is via back-projection formula, and was already presented less rigorously in the introduction. Throughout this subsection, g will be a function on the cylinder Z.

Theorem 3.16. ***Radon inversion via Fourier slices.*** *Let $f \in \mathcal{S}(\mathbb{R}^n)$. Then*

$$f(x) = \frac{1}{2} \int_{S^{n-1}} \left(\int_{-\infty}^{\infty} |\sigma|^{n-1} e^{2\pi i \sigma(x\cdot\theta)} \widehat{\mathcal{R}f}(\theta, \sigma)\, d\sigma \right) d\theta, \qquad x \in \mathbb{R}^n.$$

Proof. Set $g = \mathcal{R}f$. By inverse FT in $\mathbb{R}^n$, Thm. 3.6 and polar coordinates $\zeta = \sigma\theta$

$$\begin{aligned} f(x) &= \int_{\mathbb{R}^n} e^{2\pi i\, x\cdot\zeta} \widehat{f}(\zeta)\, d\zeta = \int_0^\infty \int_{S^{n-1}} e^{2\pi i\, x\cdot(\sigma\theta)} \widehat{f}(\sigma\theta)\, \sigma^{n-1}\, d\theta\, d\sigma \\ &= \int_{S^{n-1}} \left(\int_0^\infty e^{2\pi i \sigma(x\cdot\theta)} \widehat{g}(\theta, \sigma) \sigma^{n-1}\, d\sigma \right) d\theta = \int_{S^{n-1}} F(\theta, x)\, d\theta, \\ &\text{where } F(\theta, x) := \int_0^\infty e^{2\pi i \sigma(x\cdot\theta)} \widehat{g}(\theta, \sigma) \sigma^{n-1}\, d\sigma. \end{aligned}$$

Since the surface measure on the sphere is invariant under $\theta \mapsto -\theta$ we can write $f(x) = \frac{1}{2} \int_{S^{n-1}} \big(F(\theta, x) + F(-\theta, x)\big)\, d\theta$. To study $F(-\theta, x)$, we use Thm. 3.6 again (obtaining $\widehat{g}(-\theta, \sigma) = \widehat{f}(-\sigma\theta) = \widehat{g}(\theta, -\sigma)$) and change the variable $u = -\sigma$,

$$\begin{aligned} F(-\theta, x) &= \int_0^\infty \sigma^{n-1} e^{2\pi i\, \sigma(x\cdot(-\theta))} \widehat{g}(-\theta, \sigma)\, d\sigma = \int_0^\infty \sigma^{n-1} e^{2\pi i\, (-\sigma)(x\cdot\theta)} \widehat{g}(\theta, -\sigma)\, d\sigma \\ &= \int_{-\infty}^0 (-u)^{n-1} e^{2\pi i\, u(x\cdot\theta)} \widehat{g}(\theta, u)\, du = \int_{-\infty}^0 |u|^{n-1} e^{2\pi i\, u(x\cdot\theta)} \widehat{g}(\theta, u)\, du. \end{aligned}$$

Adding this expression to $F(\theta, x)$ gives the desired formula. □

This inversion formula was rather easy to obtain, but this simplicity comes at a price in the numerical reconstruction: it leads to Fourier-domain methods, where the data must be reorganized before applying an inverse DFT. The backprojection formula is more involved to derive, but it leads to the filtered backprojection algorithm, whose implementation is more direct.

Definition 3.17. ***Backprojection operator.*** *For $g \in C(Z)$, we define the backprojection operator $\mathcal{R}^{\#} : C(Z) \longrightarrow C(\mathbb{R}^n)$ by*

$$\mathcal{R}^{\#} g(x) := \int_{S^{n-1}} g(\theta, x \cdot \theta)\, d\theta, \qquad x \in \mathbb{R}^n.$$

As discussed in the Introduction, backprojection *spreads* each projection value back over the set of points that could have produced it. In $\mathbb{R}^3$, $\mathcal{R}f(\theta, x \cdot \theta)$ is the integral of f over the plane $\theta_{x\cdot\theta} = \{y \in \mathbb{R}^3 : \ y \cdot \theta = x \cdot \theta\}$, which passes through x and is orthogonal to θ. Thus $\mathcal{R}^{\#}\mathcal{R}f(x)$ superposes the contributions of all planes through x, producing a *blurred* reconstruction unless a filter is applied. Before deriving an inversion formula using this operator, we record some properties.

Proposition 3.18. ***Backprojection operator properties.***

1. *If $g \in C(Z)$, then $\mathcal{R}^{\#} g \in C(\mathbb{R}^n)$. And if $g \in C^1(Z)$, then $\mathcal{R}^{\#} g \in C^1(\mathbb{R}^n)$, and*

$$\frac{\partial}{\partial x_j}(\mathcal{R}^{\#} g) = \mathcal{R}^{\#}\left(\theta_j \frac{\partial g}{\partial s}\right), \qquad j = 1, \ldots, n.$$

   Finally if $g \in C^{\infty}(Z)$, then $\mathcal{R}^{\#} g \in C^{\infty}(\mathbb{R}^n)$, and for every multi-index $\alpha \in \mathbb{N}_0^n$,

$$D^{\alpha}(\mathcal{R}^{\#} g) = \mathcal{R}^{\#}\left(\theta^{\alpha} \frac{\partial^{|\alpha|} g}{\partial s^{|\alpha|}}\right).$$

2. *The backprojection and the Laplacian commute. If $\Delta = \sum_{j=1}^{n} \frac{\partial^2}{\partial x_j^2}$, and $\Box = \frac{\partial^2}{\partial s^2}$, are the usual Laplacians on $\mathbb{R}^n$, $n \geq 2$, and $\mathbb{R}$, respectively, then*

$$\mathcal{R}^{\#}(\Box g) = \Delta(\mathcal{R}^{\#} g), \qquad g \in C^{\infty}(Z).$$

3. *The operator $\mathcal{R}^{\#}$ is the real formal adjoint of $\mathcal{R}$: if $f \in \mathcal{S}(\mathbb{R}^n)$ and $g \in \mathcal{S}(Z)$, then*

$$\int_{S^{n-1}} \int_{\mathbb{R}} \mathcal{R}f(\theta, s)\, g(\theta, s)\, ds\, d\theta = \int_{\mathbb{R}^n} f(x)\, \mathcal{R}^{\#} g(x)\, dx.$$

   In particular $f * (\mathcal{R}^{\#} g) = \mathcal{R}^{\#}\big((\mathcal{R}f) *_s g\big)$.

These properties follow directly from the definition of $\mathcal{R}^{\#}$, the chain rule and Fubini's theorem; see [VG25, Sec. 3.5, pp. 22–24].

In the derivation of the Radon inversion formula via backprojection, we saw in the Introduction the appearance of the multiplier $|\sigma|$. Similarly, in higher dimensions $|\sigma|^{n-1}$ will appear. To deal with it, we need a last ingredient.

3.4.1 The Hilbert transform. $K_\varepsilon(s) := \frac{\chi_{\{|s|>\varepsilon\}}}{s}$ is locally integrable and of at most polynomial growth, hence it defines a regular tempered distribution by integration against test functions[24]. The distribution K is defined by the convergence of these actions on every Schwartz function[25] :

$$\langle K, \varphi\rangle := \lim_{\varepsilon\to 0^+} \langle K_\varepsilon, \varphi\rangle = \lim_{\varepsilon\to 0^+} \int_{|s|>\varepsilon} \frac{\varphi(s)}{s}\, ds, \qquad \varphi \in \mathcal{S}(\mathbb{R}).$$

The *Hilbert transform* is the complex-valued function $H\psi : \mathbb{R} \longrightarrow \mathbb{C}$ given by

$$H\psi(t) := \frac{i}{\pi}\,(K * \psi)(t), \qquad \psi \in \mathcal{S}(\mathbb{R}),\ t \in \mathbb{R}.$$

By the definition of convolution between a tempered distribution and a Schwartz function, we have $(K*\psi)(t) = \langle K, \psi(t-\cdot)\rangle$. Hence,

$$(K*\psi)(t) = \lim_{\varepsilon\to 0^+} \int_{|s|>\varepsilon} \frac{\psi(t-s)}{s}\, ds \implies H\psi(t) = \frac{i}{\pi} \lim_{\varepsilon\to 0^+} \int_{|s|>\varepsilon} \frac{\psi(t-s)}{s}\, ds.$$

The right-hand side is a smooth slowly increasing function of t[26]. As we saw in Preliminaries, we may also regard $H\psi$ as a tempered distribution.

The FT of K gives the distributional identity $\widehat{K} \overset{\mathcal{S}'(\mathbb{R})}{=} -i\pi\ \mathrm{sgn}$ [Cer10, Thm. 7.22, p. 203]. Therefore we obtain this identity in $\mathcal{S}'(\mathbb{R})$

$$\widehat{H\psi} = \mathcal{F}\left(\frac{i}{\pi}\, K * \psi\right) \overset{*}{=} \frac{i}{\pi}\,\widehat{K*\psi} \overset{**}{=} \frac{i}{\pi}\,\widehat{K}\,\widehat{\psi} = \frac{i}{\pi}(-i\pi\ \mathrm{sgn})\,\widehat{\psi} = \mathrm{sgn}\ \widehat{\psi}. \tag{20}$$

we have used the linearity of the FT in $*$ and the convolution rule for distributions in $**$. This is an identity in the σ-variable, i.e.

$$\left\langle \widehat{H\psi}, \varphi \right\rangle = \langle \mathrm{sgn}\ \widehat{\psi}, \varphi\rangle = \int_{\mathbb{R}} \mathrm{sgn}(\sigma)\widehat{\psi}(\sigma)\varphi(\sigma)\, d\sigma, \qquad \forall\, \varphi \in \mathcal{S}(\mathbb{R}).$$

[24] Indeed, $|K_\varepsilon(s)| \le 1/\varepsilon$. Thus it is bounded, hence of at most polynomial growth. So, $\varphi \mapsto \int_{\mathbb{R}} K_\varepsilon(s)\varphi(s)\, ds$, for $\varphi \in \mathcal{S}(\mathbb{R})$ defines a regular tempered distribution.

[25] Notice that K itself is not obtained from a locally integrable function, since $1/s$ is not locally integrable near $s = 0$.

[26] $\int_{\varepsilon<|s|<1} \frac{ds}{s} = \log(\varepsilon) - \log(\varepsilon) = 0 \implies \mathrm{p.v.}\int_{\mathbb{R}} \frac{\psi(t-s)}{s}\, ds = \int_{|s|<1} \frac{\psi(t-s)-\psi(t)}{s}\, ds + \int_{|s|\ge 1} \frac{\psi(t-s)}{s}\, ds.$

The first integral is no longer singular, because $\frac{\psi(t-s)-\psi(t)}{s} = -\int_0^1 \psi'(t-\lambda s)\, d\lambda$. The second integral is absolutely convergent, since $\psi \in \mathcal{S}(\mathbb{R})$. Moreover, the same argument applies after differentiating with respect to t: for every $k \ge 0$,

$$\frac{\partial^k}{\partial t^k}\left(\frac{\psi(t-s)-\psi(t)}{s}\right) = \frac{\psi^{(k)}(t-s)-\psi^{(k)}(t)}{s} = -\int_0^1 \psi^{(k+1)}(t-\lambda s)\, d\lambda.$$

Thus one may differentiate under the integral sign. Also, $\left|\int_{|s|\ge 1} \frac{\psi^{(k)}(t-s)}{s}\, ds\right| \le \int_{\mathbb{R}} |\psi^{(k)}(u)|\, du < \infty$, and the local part is bounded by a constant depending on $\|\psi^{(k+1)}\|_\infty$. Hence $H\psi$ is smooth, and all its derivatives are bounded; in particular, $H\psi$ is slowly increasing.

Thus the FT of $H\psi$ is computed in the distributional sense, not necessarily by an absolutely convergent Fourier integral.

Now, as promised, an inversion formula using backprojection.

Theorem 3.19. ***Radon inversion via backprojection.*** *Let $f \in \mathcal{S}(\mathbb{R}^n)$. Then*

$$f(x) = \frac{1}{2(2\pi i)^{n-1}} \mathcal{R}^{\#}\big(\Lambda(\mathcal{R}f)\big)(x), \qquad x \in \mathbb{R}^n, \tag{21}$$

where

$$\Lambda(\mathcal{R}f) := \begin{cases} \dfrac{\partial^{n-1}(\mathcal{R}f)}{\partial s^{n-1}}, & \text{if } n \text{ is odd,} \\[2ex] H\left(\dfrac{\partial^{n-1}(\mathcal{R}f)}{\partial s^{n-1}}\right), & \text{if } n \text{ is even.} \end{cases}$$

The Hilbert transform acts only on the real variable s.

Proof. Let $g = \mathcal{R}f$. We willprove, for each fixed $\theta \in S^{n-1}$, the identity

$$|\sigma|^{n-1}\widehat{g}(\theta,\sigma) = (2\pi i)^{1-n}\widehat{\Lambda g}(\theta,\sigma) \tag{22}$$

in the σ-variable. After that, We willuse the Radon inversion of Thm. 3.16.

If n is odd, then

$$\begin{aligned} |\sigma|^{n-1}\widehat{g}(\theta,\sigma) &\overset{n-1 \text{ even}}{\downarrow} \sigma^{n-1}\widehat{g}(\theta,\sigma) \overset{*}{=} (2\pi i)^{1-n}(2\pi i\sigma)^{n-1}\widehat{g}(\theta,\sigma) \overset{**}{=} (2\pi i)^{1-n}\widehat{\partial_s^{n-1}g}(\theta,\sigma) \\ &= (2\pi i)^{1-n}\widehat{\Lambda g}(\theta,\sigma), \end{aligned}$$

where in $*$ we multiplied by $\frac{(2\pi i)^{n-1}}{(2\pi i)^{n-1}}$. If n is even, then $n-1$ is odd, and hence $|\sigma|^{n-1} = \operatorname{sgn}(\sigma)\sigma^{n-1}$. Using the Hilbert transform properties we obtain, in $\mathcal{S}'(\mathbb{R})$,

$$\begin{aligned} (2\pi i)^{1-n}\widehat{\Lambda g}(\theta,\sigma) = (2\pi i)^{1-n}\mathcal{F}\left(H\left(\partial_s^{n-1}g\right)\right)(\theta,\sigma) &\overset{\text{Eq. (20)}}{\downarrow} (2\pi i)^{1-n}\operatorname{sgn}(\sigma)\widehat{\partial_s^{n-1}g}(\theta,\sigma) \\ \overset{**}{=} (2\pi i)^{1-n}\operatorname{sgn}(\sigma)(2\pi i\sigma)^{n-1}\widehat{g}(\theta,\sigma) &\underset{\substack{\uparrow \\ n-1 \text{ odd}}}{=} |\sigma|^{n-1}\widehat{g}(\theta,\sigma), \end{aligned}$$

In both equations, $**$ denotes the use of Fourier rule for derivatives

Thus (22) holds for every n. In the odd case it is an identity of functions; in the even case it is an identity of tempered distributions in the σ-variable.

Now, since $g \in \mathcal{S}(Z)$, the function $\widehat{g}(\theta,\cdot)$ is rapidly decreasing. Hence $\sigma \longmapsto |\sigma|^{n-1}\widehat{g}(\theta,\sigma)$ is integrable. Therefore its inverse FT is given by the usual integral. Applying the inverse FT in the σ-variable to equation (22) we get, for every $s \in \mathbb{R}$,

$$\mathcal{F}^{-1}\left(|\sigma|^{n-1}\widehat{g}(\theta,\sigma)\right) = \underbrace{\int_{-\infty}^{\infty} |\sigma|^{n-1} e^{2\pi i\sigma s}\widehat{g}(\theta,\sigma)\,d\sigma}_{A(s)} = \underbrace{(2\pi i)^{1-n}\Lambda g(\theta,s)}_{B(s)}. \tag{23}$$

In the even case this equality is first obtained in $\mathcal{S}'(\mathbb{R})$, that is, as an equality between the tempered distributions associated with the functions $A(s)$ and $B(s)$, but A continuous and B smooth, implies that the equality holds pointwise.

Finally, taking $s = x \cdot \theta$ in (23) and substituting into the Fourier inversion formula, we obtain

$$\begin{aligned} f(x) &\overset{\text{Thm. 3.16}}{\downarrow} \frac{1}{2}\int_{S^{n-1}} \left(\int_{-\infty}^{\infty} |\sigma|^{n-1} e^{2\pi i\, \sigma(x\cdot\theta)} \widehat{g}(\theta,\sigma)\, d\sigma \right) d\theta \\ &\overset{\text{Eq. (23)}}{\downarrow} \frac{1}{2}\int_{S^{n-1}} (2\pi i)^{1-n} \Lambda g(\theta, x\cdot\theta)\, d\theta = \frac{1}{2(2\pi i)^{n-1}} \int_{S^{n-1}} \Lambda g(\theta, x\cdot\theta)\, d\theta \\ &= \frac{1}{2(2\pi i)^{n-1}} \mathcal{R}^{\#}(\Lambda g)(x) = \frac{1}{2(2\pi i)^{n-1}} \mathcal{R}^{\#}\big(\Lambda(\mathcal{R}f)\big)(x). \end{aligned}$$

□

The backprojection inversion formula has a different character depending on the parity of n. Assume first that n is odd. Then $\Lambda g = \partial_s^{n-1} g$, and (21) becomes

$$f(x) = \frac{1}{2(2\pi i)^{n-1}} \int_{S^{n-1}} \frac{\partial^{n-1} g}{\partial s^{n-1}}(\theta, x\cdot\theta)\, d\theta.$$

For fixed θ, the value $\frac{\partial^{n-1} g}{\partial s^{n-1}}(\theta, x\cdot\theta)$ is determined only by the values of $g(\theta, t)$ for t in a neighborhood of $x \cdot \theta$. Since $g(\theta, t) = \mathcal{R}f(\theta, t)$, these values are precisely the integrals of f over the hyperplanes θ_t with t close to $x \cdot \theta$, that is, over hyperplanes parallel to $\theta_{x\cdot\theta}$ and meeting a neighborhood of x. Thus, in order to compute the integral above, we only need the values of $\mathcal{R}f$ on hyperplanes meeting a neighborhood of x. In this sense, the inversion of the Radon transform in odd dimensions is a local problem.

In even dimensions the same conclusion is not available. Indeed, in that case the operator Λ contains the Hilbert transform in the s-variable, and the Hilbert transform is not local: the value $H\psi(t)$ cannot be computed from the values of ψ on a neighborhood of t alone. In particular, in the two-dimensional case, (21) reads

$$f(x) = \frac{1}{4\pi i} \mathcal{R}^{\#}\left(H \frac{\partial g}{\partial s} \right)(x), \qquad g = \mathcal{R}f.$$

4 Discretization of the Radon transform

In the previous sections the Radon transform has been treated as a continuous object: for a given function f, the data $\mathcal{R}f(\theta, s)$ were assumed to be known for

every $(\theta, s) \in S^{n-1} \times \mathbb{R}$. This is the natural setting for the theoretical inversion formulas, but in applications one only has a finite number of samples.

Notation. For $\rho > 0$, we write $B_\rho := B(0, \rho)$, and $\chi_\rho := \chi_{B_\rho}$.

We shall think of a CT scan of a head centered at the origin and of radius at most one; thus the function that we analyze will satisfy $\operatorname{supp} f \subset \overline{B}_1$. This is a reasonable hypothesis after a suitable normalization of the field of view.

We shall also impose a restriction on the FT of f, the way in which f varies in space can be read from its FT. Assume $f \in \mathcal{S}(\mathbb{R}^n)$: Fourier inversion gives $f(x) = \int_{\mathbb{R}^n} \widehat{f}(\xi) e^{2\pi i x \cdot \xi} \, d\xi$. Thus f may be regarded as a continuous superposition of the *elementary oscillations* $x \mapsto e^{2\pi i x \cdot \xi}$, whose *frequency* is ξ. Since the component $x \mapsto e^{2\pi i x \cdot \xi}$ is periodic of period $t = \frac{1}{|\xi|}$, we can think of $|\xi|$ as the "distance" over which such exponential component may vary, i.e. the *spatial scale* of such component. Therefore, when $|\xi|$ is large, the size of $\widehat{f}(\xi)$ measures the contribution of an exponential component with rapid spatial variation. Such high-frequency components are needed to represent rapid transitions, narrow structures, or fine details accurately[27]. Small values of $|\xi|$ correspond coarser spatial behaviour.

Consequently, if one decides to work at a fixed spatial resolution $1/\Omega$, that is, to ignore variations of f occurring at spatial scales smaller than $1/\Omega$, it is natural to neglect the high-frequency part of f. Since frequencies of size $|\xi|$ describe spatial variations of scale $1/|\xi|$, this amounts to keeping only the frequency components with $|\xi| \leq \Omega$. This leads to the idealized model of bandlimited functions: one would assume that $\operatorname{supp} \widehat{f} \subset B_\Omega$.

However, exact compact support and exact bandlimitation are incompatible, except for the zero function. Indeed, as we saw in the proof of Thm. 3.12 if $f \in C_c(\mathbb{R}^n)$, then $\widehat{f}$ is real analytic. If, in addition, $\operatorname{supp} \widehat{f} \subset \overline{B}_\Omega$, then $\widehat{f}$ vanishes on the non-empty open set $\{|\xi| > \Omega\}$. By analytic continuation, $\widehat{f} \equiv 0$ on $\mathbb{R}^n$. Since the FT is injective on $L^1(\mathbb{R}^n)$, and $f \in C_c(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$, it follows that $f = 0$ a.e. Since f is continuous, $f \equiv 0$. Thus, exact localization must be replaced by approximate–localization.

The discussion above was written for Schwartz functions, where the Radon and Fourier formulas are classical. The sampling result below, however, is naturally formulated in $L^2(\mathbb{R}^n)$, since localization and approximation will be measured in terms of L^2-energy. This fits our previous setting, since $\mathcal{S}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$.

[27] This interpretation is reflected in the behaviour of derivatives under the FT. By Plancherel in $*$ and by the Fourier rule for derivatives in $**$

$$\|\nabla f\|_2^2 = \sum_{j=1}^n \left\|\partial_{x_j} f\right\|_2^2 \overset{*}{=} \sum_{j=1}^n \left\|\widehat{\partial_{x_j} f}\right\|_2^2 \overset{**}{=} \sum_{j=1}^n \int_{\mathbb{R}^n} \left|2\pi i \, \xi_j \widehat{f}(\xi)\right|^2 d\xi = (2\pi)^2 \int_{\mathbb{R}^n} |\xi|^2 \left|\widehat{f}(\xi)\right|^2 d\xi.$$

Therefore, rapid spatial variation, as measured for instance by $\|\nabla f\|_2^2$, is associated with the presence of mass of $\widehat{f}$ at large values of $|\xi|$.

Now, we introduce some vocabulary to study almost–localized functions.

Definition 4.1. *Let $f \in L^2(\mathbb{R}^n)$. We say that f is **essentially Ω-bandlimited at level** ε if*

$$\int_{B_\Omega} |\widehat{f}(\xi)|^2 \, d\xi \geq (1-\varepsilon) \, \|f\|_2^2 .$$

*We say that f is **essentially supported in B_R at level** ε if*

$$\int_{B_R} |f(x)|^2 \, dx \geq (1-\varepsilon) \, \|f\|_2^2 .$$

For $\varepsilon \in (0,1)$, We will denote by $PW^\varepsilon_{\Omega,R}$, the class of functions $f \in L^2(\mathbb{R}^n)$ which are simultaneously essentially supported in B_R and essentially Ω-bandlimited, both at level ε.

Notation. We shall call $|B_R|$ the spatial volume and $|B_\Omega|$ the frequency volume. We will set $D_{\Omega,R} := |B_R| \, |B_\Omega|$. Also, we set χ_R and χ_Ω to be the characteristic functions of B_R and B_Ω. Ω-bandlimited means *exactly* Ω-bandlimited, or essentially Ω-bandlimited at level 0. For $\varepsilon > 0$, we shall loosely refer to the corresponding notions as *almost* space (or band) localization.

These definitions are energy conditions. For instance, $\frac{\int_{B_R} |f(x)|^2 \, dx}{\|f\|_2^2}$ is the fraction of the L^2-energy of f contained in B_R. Thus f is essentially supported in B_R at level ε precisely when the energy outside B_R is at most an ε-fraction of the total energy: $\int_{\mathbb{R}^d \setminus B_R} |f(x)|^2 \, dx \leq \varepsilon \, \|f\|_2^2$. The same interpretation applies to the FT in the definition of essential band limitation.

Our goal is to estimate these almost–localized functions. A theorem of Landau, Slepian and Pollak does precisely this.

Theorem 4.2. *There exist finite-dimensional subspaces $H_{\Omega,R} \subset L^2(\mathbb{R}^n)$ such that*

$$\lim_{R\Omega \to \infty} \frac{\dim H_{\Omega,R}}{D_{\Omega,R}} = 1.$$

Moreover,

$$\lim_{\varepsilon \to 0^+} \limsup_{R\Omega \to \infty} \sup_{f \in PW^\varepsilon_{\Omega,R}, f \neq 0} \frac{\operatorname{dist}(f, H_{\Omega,R})}{\|f\|_2} = 0.$$

This order of dimension is asymptotically optimal: there are no other subspaces $h_{\Omega,R} \subset L^2(\mathbb{R}^n)$ with the same approximation property and such that $\liminf_{R\Omega \to \infty} \frac{\dim h_{\Omega,R}}{\dim H_{\Omega,R}} < 1$.

Let us first interpret the statement before proving it. A common heuristic in *sampling theory* says that a space and band localized function can be approximated by a finite number of parameters, which is governed by the product of the spatial and frequency volumes. In our setting, functions in $PW^\varepsilon_{\Omega,R}$ are *essentially* localized

in the spatial ball B_R and in the frequency ball B_Ω. The heuristic suggests the quantity

$$|B_1|R^n\,|B_1|\Omega^n = |B_1|^2\,(R\Omega)^n = D^n_{R,\Omega}.$$

Thm. 4.2 makes this heuristic precise: it states that, although $PW^\varepsilon_{\Omega,R}$ is not finite-dimensional, its elements can be approximated, uniformly as $R\Omega \to \infty$, by elements of a subspace $H_{\Omega,R} \subset L^2(\mathbb{R}^n)$.

In the present problem the scalar data are not point values of f, but values of its Radon transform, $\mathcal{R}f(\theta_j, s_l)$. Each such datum gives one linear condition on the approximant. Indeed, if $\{\phi_1, \dots, \phi_N\}$ is a basis of $H_{\Omega,R}$ and $h = \sum_{k=1}^N c_k\phi_k$ then, by linearity of the Radon transform,

$$\mathcal{R}h(\theta_j, s_l) = \int_{(\theta_j)_{s_l}} h(x)\,dx = \int_{(\theta_j)_{s_l}} \sum_{k=1}^N c_k\phi_k(x)\,dx = \sum_{k=1}^N c_k\,\mathcal{R}\phi_k(\theta_j, s_l).$$

Thus the Radon datum $\mathcal{R}f(\theta_j, s_l)$ produces the linear equation $\sum_{k=1}^N c_k\,\mathcal{R}\phi_k(\theta_j, s_l) = \mathcal{R}f(\theta_j, s_l)$. If we take M samples of the Radon transform at the parameter values $(\theta_1, s_1), \dots, (\theta_M, s_M)$, we obtain the linear system

$$\begin{pmatrix} \mathcal{R}\phi_1(\theta_1, s_1) & \cdots & \mathcal{R}\phi_N(\theta_1, s_1) \\ \vdots & \ddots & \vdots \\ \mathcal{R}\phi_1(\theta_M, s_M) & \cdots & \mathcal{R}\phi_N(\theta_M, s_M) \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_N \end{pmatrix} = \begin{pmatrix} \mathcal{R}f(\theta_1, s_1) \\ \vdots \\ \mathcal{R}f(\theta_M, s_M) \end{pmatrix}.$$

Thus, each Radon datum gives one scalar linear equation for the coefficients of an approximant in $H_{\Omega,R}$. Since $\dim H_{\Omega,R} \sim D_{\Omega,R}$, the theorem suggests that one should not expect to need more than about $D_{\Omega,R}$ independent scalar measurements in order to describe functions in $PW^\varepsilon_{\Omega,R}$ up to the prescribed resolution. However, the concrete choice of Radon samples and the stability of the resulting linear system are separate numerical issues.

Increasing R means covering a larger spatial region, while increasing Ω means resolving finer details. Both increase the number of non-redundant data required by this finite-dimensional approximation. In dimension n, this number grows like $(R\Omega)^n$.[28]

Idea of the proof of Thm. 4.2: First we would like to define some finite-dimensional subspace of $L^2(\mathbb{R}^n)$. To do that, a simple idea is to try to "project" functions in $PW^\varepsilon_{\Omega,R}$ to some subspace of PW_Ω that are also essentially B_R–supported. A simple approach is to just use the characteristic functions χ_R and χ_Ω. Define the

[28] it is worth noting that Thm. 4.2 is a sufficiency statement. It does not say that a number of data of order $(R\Omega)^n$ is always necessary, nor does it exclude that particular functions, or classes with additional structure, may be approximated with fewer data. Rather, it says that a number of suitably independent Radon data of this order is enough to guarantee a uniform L^2-approximation of any function essentially localized in space and in frequency.

concentration operator $C_{\Omega,R}$ as

$$C_{\Omega,R} : PW_\Omega \longrightarrow PW_\Omega, \qquad C_{\Omega,R} f := P_\Omega(\chi_R \cdot f), \qquad P_\Omega f := \mathcal{F}^{-1}\big(\chi_\Omega \cdot \mathcal{F}(f)\big).$$

From now and on, We will omit the $\cdot$ sign. $C_{\Omega,R}$ first cuts the function in space, keeping only its part inside B_R, and then projects back onto the bandlimited space PW_Ω. We'll find that this operator gives an orthonormal basis $(f_k)_{k\in\mathbb{N}}$; and that it has finitely many eigenvalues close to one, while the rest tend to zero (Prop. 4.3). This gives us the idea of defining $H_{\Omega,R}$ as the span of the first N eigenfunctions.

If $f \in PW_\Omega$ is essentially B_R–supported, We willsee that its expansion in $(f_k)_{k\in\mathbb{N}}$ will not have much mass in the directions with small eigenvalues, because those directions do not contribute to the mass inside B_R. Hence the tail outside $H_{\Omega,R}$ will be small. Then, for $f \in PW^\varepsilon_{\Omega,R}$, we first project it in frequency onto PW_Ω, control the error by Plancherel, and apply the previous case.

We shall use the following spectral properties of the concentration operator, taken from [CMOCP07, Sec. 3, Props. 18–19]. They are stated without proof in the main text, since the proof is interesting but too long to include here. For completeness, an expanded proof is given in Appendix 1, although it is not needed for the rest of the work.

Proposition 4.3. *The operator $C_{\Omega,R}$ is linear, bounded, compact, self-adjoint, positive, injective and satisfies $\langle C_{\Omega,R} f, f\rangle = \int_{B_R} |f(x)|^2\,dx$.*

If $\lambda_k(\Omega,R)$ for $k \in \mathbb{N}$ are its eigenvalues, arranged in decreasing order, then $\lambda_k(\Omega,R) \in (0,1)$ and, for any threshold $t \in (0,1)$

$$D_{\Omega,R} - \frac{o(D_{\Omega,R})}{1-t} \le \#\{k : \lambda_k(\Omega,R) > t\} \le D_{\Omega,R} + \frac{o(D_{\Omega,R})}{t}, \qquad R\Omega \to \infty.$$

Observe that for every fixed t, the number of eigenvalues *above* it is asymptotically equal to $D_{\Omega,R}$. In particular, $\#\{k : t < \lambda_k(\Omega,R) < 1-t\} = o(D_{\Omega,R})$.[29] Thus, up to a negligible number of eigenvalues, the spectrum is concentrated near 0 and near 1. The eigenvalues close to 1 correspond to bandlimited functions which are well concentrated in B_R, whereas the eigenvalues close to 0 correspond to bandlimited functions whose mass is mostly outside B_R.

Proof. (*Thm. 4.2*). Set $\lambda_k = \lambda_k(\Omega,R)$ and $C := C_{\Omega,R}$ We first choose two cut-off

[29] Indeed, if $t \ge 1/2$, then the set is empty. If $0 < t < 1/2$, then
$\#\{k : t < \lambda_k < 1-t\} \le \#\{k : \lambda_k > t\} - \#\{k : \lambda_k > 1-t\} \overset{*}{\le} \left(D_{\Omega,R} + \frac{o(D_{\Omega,R})}{t}\right) - \left(D_{\Omega,R} - \frac{o(D_{\Omega,R})}{t}\right)$
where in the * we applied the estimate first with t, and then with $1-t$.

indices. By Prop. 4.3, for each fixed $j \geq 2$,

$$\begin{aligned}
&\#\left\{k : \lambda_k > \frac{1}{j}\right\} \leq D_{\Omega,R} + j\,o(D_{\Omega,R}) \overset{*}{\leq} D_{\Omega,R} + \frac{1}{j} D_{\Omega,R} = \left(1 + \frac{1}{j}\right) D_{\Omega,R}, \\
&\#\left\{k : \lambda_k > 1 - \frac{1}{j}\right\} \geq D_{\Omega,R} - \frac{o(D_{\Omega,R})}{1 - \left(1 - \frac{1}{j}\right)} = D_{\Omega,R} - j\,o(D_{\Omega,R}) \underset{*}{\geq} \left(1 - \frac{1}{j}\right) D_{\Omega,R}.
\end{aligned} \tag{24}$$

where $(*)$ holds for each fixed j and $R\Omega$ sufficiently large. Choosing $j = j(\Omega, R) \to \infty$ slowly enough, these inequalities still hold as $R\Omega \to \infty$. Define

$$n(\Omega R) := \left\lfloor \left(1 - \frac{1}{j(\Omega,R)}\right) D_{\Omega,R} \right\rfloor, \qquad N(\Omega R) := \left\lfloor \left(1 + \frac{1}{j(\Omega,R)}\right) D_{\Omega,R} \right\rfloor + 1.$$

Since the eigenvalues are arranged decreasingly and by (24), the first $n(\Omega R)$ of them are asymptotically close to 1, while the eigenvalues from $N(\Omega R)$ onward are asymptotically close to 0. So, we can write $\lambda_{n(\Omega R)}(\Omega, R) \xrightarrow{R\Omega\nearrow\infty} 1$, and $\lambda_{N(\Omega R)}(\Omega, R) \xrightarrow{R\Omega\nearrow\infty} 0$. Also,

$$\frac{N(\Omega R)}{D_{\Omega,R}} \xrightarrow{R\Omega\nearrow\infty} 1, \qquad \frac{n(\Omega R)}{N(\Omega R)} \xrightarrow{R\Omega\nearrow\infty} 1.$$

Now set $H_{\Omega,R} := \operatorname{span}\{f_k : 1 \leq k < N(\Omega R)\}$. Then $\frac{\dim H_{\Omega,R}}{D_{\Omega,R}} = \frac{N(\Omega R)-1}{D_{\Omega,R}} \to 1$.

Let $f \in PW_\Omega$, $\|f\|_2 = 1$, and suppose that f is essentially supported in B_R at level ε. Write $f = \sum_{k=1}^{\infty} v_k f_k$, with $v_k = \langle f, f_k\rangle$ (thus $\sum_{k=1}^{\infty} |v_k|^2 = 1$). Then, on one hand, for $R\Omega$ large enough so that $\lambda_{N(\Omega R)} \leq \varepsilon$,

$$\begin{aligned}
1 - \varepsilon \leq \int_{B_R} |f(x)|^2\,dx = \langle C_{\Omega,R} f, f\rangle &= \sum_{k=1}^{\infty} \lambda_k |v_k|^2 \\
&\leq \sum_{k<N(\Omega R)} |v_k|^2 + \lambda_{N(\Omega R)} \sum_{k\geq N(\Omega R)} |v_k|^2 \leq \sum_{k<N(\Omega R)} |v_k|^2 + \varepsilon.
\end{aligned}$$

Hence $\sum_{k\geq N(\Omega R)} |v_k|^2 \leq 2\varepsilon$. Now, on the other hand, let P_H denote the orthogonal projection onto $H_{\Omega,R}$. We have $P_H f = \sum_{k<N} v_k f_k$, and $f - P_H f = \sum_{k\geq N} v_k f_k$. Therefore, by Parseval,

$$\operatorname{dist}(f, H_{\Omega,R})^2 = \|f - P_H f\|_2^2 = \sum_{k\geq N(\Omega R)} |v_k|^2 \leq 2\varepsilon.$$

We pass now to the class $PW^{\varepsilon}_{\Omega,R}$, with $\varepsilon \in (0,1)$. Let $f \in PW^{\varepsilon}_{\Omega,R}$, $\|f\|_2 = 1$, and put $u := P_\Omega f = \mathcal{F}^{-1}\big(\chi_\Omega \cdot \mathcal{F}(f)\big)$. Then $u \in PW_\Omega$, and by Plancherel,

$$\|f - u\|_2^2 = \left\|\widehat{f - u}\right\|_2^2 = \int_{\mathbb{R}^n \setminus B_\Omega} |\widehat{f}(\xi)|^2\,d\xi \leq \varepsilon; \qquad \|u\|_2^2 \geq 1 - \varepsilon.$$

Moreover, $\|(1-\chi_R)u\|_2 \le 2\sqrt{\varepsilon}$.[30] Thus

$$\int_{\mathbb{R}^n\setminus B_R}\left|\frac{u(x)}{\|u\|_2}\right|^2 dx = \frac{\|(1-\chi_R)u\|_2^2}{\|u\|_2^2} =\le \frac{4\varepsilon}{1-\varepsilon} =: \alpha_\varepsilon.$$

Applying the previous estimate to $u/\|u\|_2$, we get $\operatorname{dist}(u, H_{\Omega,R}) \le \sqrt{2\alpha_\varepsilon}$. Finally,

$$\operatorname{dist}(f, H_{\Omega,R}) \le \|f-u\|_2 + \operatorname{dist}(u, H_{\Omega,R}) \le \sqrt{\varepsilon} + \sqrt{\frac{8\varepsilon}{1-\varepsilon}} \xrightarrow{\varepsilon\searrow 0} 0,$$

which proves the approximation statement.

It remains to justify the optimality of the order of the dimension. Define $G_{\Omega,R} := \operatorname{span}\{f_k : 1 \le k < n(\Omega R)\}$. Then $\frac{\dim(G_{\Omega,R})}{\dim(H_{\Omega,R})} = \frac{n(\Omega R)-1}{N(\Omega R)-1} \xrightarrow{R\Omega\nearrow\infty} 1$.

If $f \in G_{\Omega,R}$ and $\|f\|_2 = 1$, write again $f = \sum_{k<n(\Omega R)} v_k f_k$, with $v_k = \langle f, f_k\rangle$ (thus $\sum_{k<n(\Omega R)} |v_k|^2 = 1$). Then

$$\begin{aligned}\int_{B_R} |f(x)|^2\,dx &\overset{\text{Prop. 4.3}}{=} \langle C_{\Omega,R} f, f\rangle = \sum_{k<n(\Omega R)} \lambda_k |v_k|^2 \\ &\ge \lambda_{n(\Omega R)} \sum_{k<n(\Omega R)} |v_k|^2 = \lambda_{n(\Omega R)} \xrightarrow{R\Omega\nearrow\infty} 1.\end{aligned}$$

Therefore, for every fixed $\varepsilon > 0$, all unit vectors of $G_{\Omega,R}$ belong to $PW^\varepsilon_{\Omega,R}$ when $R\Omega$ is large enough.

Now let $E_{\Omega,R} \subset L^2(\mathbb{R}^n)$ be any subspace which approximates all unit vectors in $PW^\varepsilon_{\Omega,R}$ with relative error strictly smaller than 1. If $\dim E_{\Omega,R} < \dim G_{\Omega,R}$, then $G_{\Omega,R} \cap E^\perp_{\Omega,R}$ contains a non-zero vector g. After normalizing it, we would have

$$\operatorname{dist}(g, E_{\Omega,R}) = \left\|g - P_{E_{\Omega,R}} g\right\|_2 = \|g\|_2 = 1,$$

contradicting the assumed error bound. Hence $\dim E_{\Omega,R} \ge \dim G_{\Omega,R}$. Since $\dim G_{\Omega,R}/\dim H_{\Omega,R} \to 1$, the dimension of $H_{\Omega,R}$ is asymptotically optimal. □

4.1 The semidiscrete case

The previous section gave an abstract estimate for the number of scalar parameters needed to approximate any almost–localized function. We now specialize these considerations to Radon data for such functions and ask how many directions are sufficient, assuming that the radial variable remains continuous.

More precisely, suppose that for a finite set $\Theta \subset S^{n-1}$ the data $\mathcal{R}f(\theta, s)$, $\theta \in \Theta$, $s \in \mathbb{R}$ is known. Since $\widehat{\mathcal{R}f}(\theta, \sigma) = \widehat{f}(\sigma\theta)$, therefore, choosing finitely many directions is equivalent to observing $\widehat{f}$ only on finitely many frequency lines through

[30] $\|(1-\chi_R)u\|_2 = \|(1-\chi_R)f + (1-\chi_R)(u-f)\|_2 \le \|(1-\chi_R)f\|_2 + \|u-f\|_2 \le 2\sqrt{\varepsilon}$.

the origin. Our problem is to understand how much of that f ($\mathcal{R}f$) is determined by finitely many such frequency lines (directions).

For each fixed direction θ, set $g_\theta(\sigma) := \widehat{\mathcal{R}f}(\theta, \sigma)$. On one hand

$$\mathcal{F}_\sigma g_\theta(s) = \mathcal{F}_\sigma \left(\widehat{\mathcal{R}f}(\theta, \cdot)\right)(s) = \mathcal{F}^2\left(\mathcal{R}f(\theta, \cdot)\right)(s) = \mathcal{R}f(\theta, -s). \tag{25}$$

Therefore, if $f \in C_c^\infty(B_1)$, then $\mathcal{R}f(\theta, \cdot)$ is supported in $[-1, 1]$, and hence $\operatorname{supp} \widehat{g}_\theta \subset [-1, 1]$, i.e. g_θ is bandlimited.

On the other hand, recall that by Thm. 3.14, the moments of the Radon data are functions only of θ, and they are also are linked to the frequency slice:

$$\frac{d^j}{d\sigma^j} g_\theta(\sigma)\Big|_{\sigma=0} = \frac{d^j}{d\sigma^j} \widehat{\mathcal{R}f}(\theta, \sigma)\Big|_{\sigma=0} = (-2\pi i)^j \int_{\mathbb{R}} s^j \mathcal{R}f(\theta, s)\, ds = \mu_j(\theta). \tag{26}$$

Thus, if μ_j vanish, for $j = 0, \ldots, k$, then $g_\theta(0) = g'_\theta(0) = \cdots = g_\theta^{(k)}(0) = 0$.

The following lemma, which we state without proof (see [CMOCP07, Sec. 3.1, Lem. 20]), relates both facts: a one-dimensional bandlimited function with many vanishing derivatives at the origin cannot keep most of its weighted L^2-mass near the origin.

Lemma 4.4. *Let $\alpha \in \mathbb{R}$. For every $0 < \delta < 1$, there is an integer $k_0 = k_0(\alpha, \delta) \geq 0$ such that, if $k \geq k_0$, $\operatorname{supp} \widehat{g} \subset [-1, 1]$, and $g^{(j)}(0) = 0$, $j = 0, \ldots, k$, then*

$$\int_{-(1-\delta)k/2}^{(1-\delta)k/2} |g(x)|^2 (1 + |x|)^\alpha \, dx \leq \delta \int_{\mathbb{R}} |g(x)|^2 (1 + |x|)^\alpha \, dx,$$

provided that the integral on the right is finite.

Recall that Thm. 3.14 guarantees μ_m is a homogeneous polynomial of degree m restricted to the sphere. Such polynomials can be decomposed into "spherical harmonics", which will allow us to decide when finitely many directions are enough to recover the relevant moment information.

A polynomial q such that $\Delta q = 0$, where $\Delta = \sum_{j=1}^n \partial_{x_j}^2$, is called a **harmonic polynomial**. A polynomial $h_k \in \mathcal{P}_k^{\text{hom}}(\mathbb{R}^n)$ that is harmonic is thus a **homogeneous harmonic polynomial of degree** k. We denote by $\mathcal{H}_k(\mathbb{R}^n)$ the complex vector **space of harmonic homogeneous polynomials of degree** k, that is, $\mathcal{H}_k(\mathbb{R}^n) := \{h_k \in \mathcal{P}_k^{\text{hom}}(\mathbb{R}^n) : \Delta h_k = 0\}$.

Every harmonic polynomial q can be written $q = h_0 + h_1 + \ldots$, with $h_k \in \mathcal{H}_k(\mathbb{R}^n)$. Similarly, for any $p_k \in \mathcal{P}_k^{\text{hom}}(\mathbb{R}^n)$, there are unique $h_{k-2j} \in \mathcal{H}_{k-2j}(\mathbb{R}^n)$ such that (see [ABR01, p. 75; Thm. 5.7, p. 77])

$$p_k(x) = \sum_{j=0}^{\lfloor k/2 \rfloor} |x|^{2j} h_{k-2j}(x) = h_k(x) + |x|^2 h_{k-2}(x) + |x|^4 h_{k-4}(x) + \cdots. \tag{27}$$

The restrictions to S^{n-1} of the elements of $\mathcal{H}_\ell(\mathbb{R}^n)$ are called **spherical harmonics of degree** ℓ. We denote their complex vector space by $\mathcal{H}_\ell(S^{n-1})$, that is, $\mathcal{H}_\ell(S^{n-1}) := \{h_{\ell|S^{n-1}} : h_\ell \in \mathcal{H}_\ell(\mathbb{R}^n)\}$.[31]

After restriction to the unit sphere, the powers of $|x|^2$ in (27) disappear: if $\theta \in S^{n-1}$, then $|\theta| = 1$, and hence $p_k(\theta) = h_k(\theta) + h_{k-2}(\theta) + h_{k-4}(\theta) + \cdots$. Thus $p_{k|S^{n-1}}$ coincides on S^{n-1} with the harmonic polynomial on $\mathbb{R}^n$ $H_k := h_k + h_{k-2} + h_{k-4} + \cdots$, which has degree at most k and the same parity as k. Equivalently, $p_{k|S^{n-1}}$ is a finite sum of spherical harmonics $h_{k-2j|S^{n-1}}$ of degrees $k, k-2, k-4, \ldots$.[32] In particular, Thm. 3.14 says that μ_m is the restriction to S^{n-1} of some $p_m \in \mathcal{P}_m^{\hom}(\mathbb{R}^n)$, so

$$\begin{aligned}\mu_m &\overset{\text{Thm. 3.14}}{\downarrow} = p_{m|S^{n-1}} \overset{\text{Eq. (27)}}{\downarrow} = \left(h_m + |x|^2 h_{m-2} + + \cdots\right)_{|S^{n-1}} \\ &= h_{m|S^{n-1}} + h_{m-2|S^{n-1}} + \cdots = (h_m + h_{m-2} + + \cdots)_{|S^{n-1}} = H_{m|S^{n-1}}.\end{aligned}$$

Notice $H_m(-x) = h_m(-x) + h_{m-2}(-x) + \cdots = (-1)^m h_m(x) + (-1)^{m-2} h_{m-2}(x) + \ldots$, thus H_m is a harmonic polynomial that is even if and only if m is even and odd if and only if m is odd. Now set $H'_m(\mathbb{R}^n)$ as the *set of harmonic polynomials of degree at most m that satisfy this degree–parity matching*. Observe that if $k \leq m$, $H'_k \subset H'_m$. With this naming, μ_m coincides with an element of H'_m on S^{n-1}.

Definition 4.5. *A set $\Theta_m \subset S^{n-1}$ is m-**resolvent** if it is a uniqueness set for H'_m. Equivalently, if $h \in H'_m$ satisfies $h(\theta) = 0$ for any $\theta \in \Theta_m$ implies that $h \equiv 0$ on S^{n-1}.*[33]

Observe that $|\Theta_m| \geq \dim H'_m$. Suppose that $\Theta_m = \{\theta_1, \ldots, \theta_N\}$, then the evaluation map $H'_m \longrightarrow \mathbb{C}^N$, such that $h \mapsto \big(h(\theta_1), \ldots, h(\theta_N)\big)$ must be injective. Indeed, if this map had a non-trivial kernel, there would exist $0 \neq h \in H'_m$ such that $h(\theta_1) = \cdots = h(\theta_N) = 0$, contradicting the fact that Θ_m is m-resolvent.

Theorem 4.6. ***Semidiscrete estimate.*** *Let $0 < \varepsilon < 1$, $m \in \mathbb{N}$, $\Theta_m \subset S^{n-1}$ be m-resolvent, set $a_m := \frac{(1-\varepsilon)m}{2}$ and $f \in C_c^\infty(B_1)$. If $\mathcal{R}f(\theta, s) = 0$ for $\theta \in \Theta_m$, $s \in \mathbb{R}$ then*

$$\int_{B_{(1-\varepsilon)m/2}} |\widehat{f}(\zeta)|^2 \, d\zeta \leq C_m \, \|f\|_2^2 ,$$

[31] With respect to the usual inner product $\langle u, v\rangle_{S^{n-1}} := \int_{S^{n-1}} u(\theta)\overline{v(\theta)}\, d\sigma(\theta)$, the spaces $\mathcal{H}_\ell(S^{n-1})$ are mutually orthogonal, and span $L^2(S^{n-1})$ in the Hilbert-space sense [ABR01, Thm. 5.12, pp. 80–81].

[32] Observe that $h_{\ell|S^{n-1}}$ is a *function on the sphere*. We call it "harmonic" because it comes from a *homogeneous harmonic* polynomial. However, $h_\ell|_{S^{n-1}}$ is not homogeneous in itself, nor does it have an intrinsic degree in the usual polynomial sense, since both notions are inherited from the homogeneous polynomial h_ℓ defined on $\mathbb{R}^n$. For instance, $p_2(x) = |x|^2$ is homogeneous of degree 2, but $p_2|_{S^{n-1}} = 1$, which is a spherical harmonic of degree 0.

[33] As m increases, the space H'_m becomes larger and so the condition of being m-resolvent becomes more demanding. Indeed, since $H'_k \subset H'_m$ for every $k \leq m$, every m-resolvent set is automatically k-resolvent for every $k \leq m$.

where $C_m > 0$ is independent of f and $C_m \to 0$ as $m \to \infty$.

Proof. If $\mathcal{R}f(\theta, s) \equiv 0$ in Θ, then by the identities in (26) we have $\mu_k \equiv 0$ on Θ for all $k \in \mathbb{N} \cup \{0\}$. Meanwhile, our previous discussion showed that $\mu_k \in H'_m$. Since Θ_m is m-resolvent and μ_k vanishes on Θ_m, we obtain $\mu_k \equiv 0$ on S^{n-1} for $k = 0, \dots, m$.

Now fix $\theta \in S^{n-1}$. $\mu_k(\theta) = 0$ for $k = 0, \dots, m$ is equivalent, because of (26), to $\partial_\sigma^j \widehat{\mathcal{R}f}(\theta, \sigma)\Big|_{\sigma=0} = 0$ for $k = 0, \dots, m$. As we did before, set $g_\theta(\sigma) := \widehat{\mathcal{R}f}(\theta, \sigma)$. The hypothesis $\operatorname{supp} f \subset B(0,1)$ implies, as we saw below (25), that $\operatorname{supp} \widehat{g}_\theta \subset [-1, 1]$. It follows that Lem. 4.4 applies to g_θ. With the notation of this theorem, we apply Lem. 4.4 with parameters $\alpha = n - 1$ and $k = m$. Thus, for a suitable constant $C_m \to 0$, we obtain

$$\begin{aligned} \int_{-a_m}^{a_m} |g_\theta(\sigma)|^2 |\sigma|^{n-1} \, d\sigma &\le \int_{-a_m}^{a_m} |g_\theta(\sigma)|^2 (1 + |\sigma|)^{n-1} \, d\sigma \\ &\overset{\text{Thm. 4.4}}{\downarrow} \\ &\le C_m \int_{\mathbb{R}} |g_\theta(\sigma)|^2 (1 + |\sigma|)^{n-1} \, d\sigma. \end{aligned} \tag{28}$$

We now integrate (28) with respect to $\theta \in S^{n-1}$. By Thm. 3.6, $g_\theta(\sigma) = \widehat{f}(\sigma\theta)$, and by Tonelli and polar coordinates, the left-hand side integrated over S^{n-1} of (28) becomes

$$\int_{S^{n-1}} \int_{-a_m}^{a_m} |\widehat{f}(\sigma\theta)|^2 |\sigma|^{n-1} \, d\sigma \, d\theta = 2 \int_{B(0,a_m)} |\widehat{f}(\zeta)|^2 \, d\zeta.$$

On the other hand, the integral over S^{n-1} of the right-hand side of (28) can be split into two integrals: $\int_{S^{n-1}} \int_{|\sigma|>1}$ and $\int_{S^{n-1}} \int_{|\sigma|\le 1}$. First $\int_{S^{n-1}} \int_{|\sigma|>1}$

$$\begin{aligned} \int_{S^{n-1}} \int_{|\sigma|\ge 1} |\widehat{f}(\sigma\theta)|^2 (1 + |\sigma|)^{n-1} \, d\sigma \, d\theta &\le \int_{S^{n-1}} \int_{|\sigma|\ge 1} |\widehat{f}(\sigma\theta)|^2 (2\,|\sigma|)^{n-1} \, d\sigma \, d\theta \\ &\overset{\xi=\sigma\theta}{\downarrow} \\ &\le 2^{n-1} \int_{\mathbb{R}^n} |\widehat{f}(\zeta)|^2 \, d\zeta \overset{\text{Plancherel}}{=} 2^{n-1} \|f\|_2^2, \end{aligned}$$

For $|\sigma| \le 1$, using $(1 + |\sigma|)^{n-1} \le 2^{n-1}$ and $|\widehat{f}(\sigma\theta)| \le \|f\|_1 \le |B(0,1)|^{1/2} \|f\|_2$, we get

$$\int_{S^{n-1}} \int_{|\sigma|\le 1} |\widehat{f}(\sigma\theta)|^2 (1 + |\sigma|)^{n-1} \, d\sigma \, d\theta \le 2^n |S^{n-1}| \, |B(0,1)| \, \|f\|_2^2.$$

Finally we absorb the dimensional constants into C_m. □

The main application of the theorem is the following corollary that, roughly speaking, says that an $m/2$-bandlimited function is determined by its Radon transform on a set of directions that is m-resolvent.

Corollary 4.7. *Assume the hypothesis of 4.6, set r such that $0 < r < 1 - C_m$ and assume that f is essentially a_m bandlimited with level r. If we know $\mathcal{R}f(\theta, s)$ for every $\theta \in \Theta_m$ and all $s \in \mathbb{R}$, then we know $\mathcal{R}f(\theta, s)$ for every $\theta \in S^{n-1}$ and all $s \in \mathbb{R}$.*

Proof. Suppose that there is another admissible function g, also essentially a_m-bandlimited at level r, and $\mathcal{R}g(\theta, s) = \mathcal{R}f(\theta, s)$ for all $\theta \in \Theta_m$, $s \in \mathbb{R}$. Set $h := f - g$. Then, by linearity, $\mathcal{R}h(\theta, s) = 0$, $\forall \theta \in \Theta_m$, $s \in \mathbb{R}$. Moreover, h is essentially a_m-bandlimited at level r. We have

$$\int_{B_{a_m}} |\widehat{h}(\xi)|^2 \, d\xi \overset{\text{Thm. 4.6}}{\leq} C_m \left\|\widehat{h}\right\|_2^2; \qquad \int_{\mathbb{R}^n \setminus B_{a_m}} |\widehat{h}(\xi)|^2 \, d\xi \overset{a_m \text{ bandlimited with level } r}{\leq} r \left\|\widehat{h}\right\|_2^2.$$

Therefore

$$\left\|\widehat{h}\right\|_2^2 = \int_{B_{a_m}} |\widehat{h}(\xi)|^2 \, d\xi + \int_{\mathbb{R}^n \setminus B_{a_m}} |\widehat{h}(\xi)|^2 \, d\xi \leq C_m \left\|\widehat{h}\right\|_2^2 + r \left\|\widehat{h}\right\|_2^2 = (C_m + r) \left\|\widehat{h}\right\|_2^2.$$

Since $r < 1 - C_m$, we have $C_m + r < 1$. Hence the previous inequality forces $\left\|\widehat{h}\right\|_2 = 0$. By Plancherel, $h = 0$ in $L^2(\mathbb{R}^n)$. Thus $f = g$. Hence the data on $\Theta_m \times \mathbb{R}$ determine f, and consequently determine $\mathcal{R}f(\theta, s)$ for every $\theta \in S^{n-1}$ and every $s \in \mathbb{R}$. □

Cor. 4.7. requires the exact knowledge of the $\mathcal{R}f(\theta, s)$ in all the directions of Θ_m. It would be desirable a similar result with only an "approximate" knowledge. We now estimate how many discrete values of s would be needed in order to replace this continuous knowledge by a finite number of samples.

On one hand, under the hypotheses of Cor. 4.7, we have that f being essentially bandlimited with band a_m with level r, means that $\widehat{f}$ is essentially supported in B_{a_m} with level r. This implies that $\widehat{f}$ is essentially supported in $B_{m/2}$ with level r. This corresponds, by Thm. 3.6, to $\widehat{\mathcal{R}f}(\theta, \cdot)$ being essentially supported in $[-m/2, m/2]$ to some level $\widetilde{r}_\theta$ ($\widetilde{r}$ might depend on θ since we are restricting $\widehat{f}$ to the line $\{\sigma\theta : \sigma \in \mathbb{R}\}$). Thus, $\mathcal{R}f(\theta, s)$ is essentially $m/2$ bandlimited with level $\widetilde{r}_\theta$.

On the other hand, since $\operatorname{supp}(f) \subset \overline{B_1}$, we have that for each fixed θ, $\mathcal{R}f(\theta, \cdot)$ is a one-dimensional function supported in $[-1, 1]$, and hence also essentially supported in $[-1, 1]$ for every level of concentration, in particular for $\widetilde{r}_\theta$

Thus, for that fixed θ, $\mathcal{R}f(\theta, \cdot) \in PW^{\widetilde{r}_\theta}_{m/2;1}$ and, by Thm. 4.2 in dimension 1, the number of samples needed to approximate $\mathcal{R}f(\theta, \cdot)$ from a subspace of $L^2(\mathbb{R})$ is

$$\operatorname{vol}([-1, 1]) \operatorname{vol}([-m/2, m/2]) = 2 \cdot m = 2m.$$

Since this count has to be repeated for every direction in Θ_m, the total number of scalar Radon samples is, in first order (see [CMOCP07, Sec. 2.5, pp. 14–16]),

$$2m \operatorname{card}(\Theta_m) \geq 2m \dim(H'_m) = 2m \cdot 2 \binom{m + n - 1}{n - 1} = 2m \cdot 2 \frac{m^{n-1}}{(n-1)!} \left(1 + o(1/m)\right).$$

Therefore, we need a minimum of $\frac{4m^n}{(n-1)!}$ samples to apply the Cor. 4.7.

A different estimate of the number of needed samples according to Thm. 4.2 is

$$\operatorname{vol}(B(0,m/2))\operatorname{vol}(B(0,1)).$$

In particular, for $n=2$, the Thm. 4.2 count gives $\operatorname{vol}(B(0,m/2))\operatorname{vol}(B(0,1)) \simeq 2.467m^2$ whereas the semidiscrete count gives $4m^2$. Thus, in dimension 2, our sufficient condition is larger by the factor $\frac{4}{\pi^2/4} = \frac{16}{\pi^2} \simeq 1.62$. It seems that the sufficient condition that we get is of a factor $4/2.467$ redundant, therefore there is some room for improvement.

5 Reconstruction algorithms

Given finitely many Radon data, we want to construct an approximation of f. We describe two methods.

For simplicity, we work in the plane. Let f be supported in $\mathbb{D} = B(0,1)$, extended by zero outside $\mathbb{D}$, and assume that $f \in L^2(\mathbb{D})$, assume that $\mathcal{R}f(\theta,s)$ is well-defined for every $(\theta,s) \in S^1 \times \mathbb{R}$, and that f is essentially Ω-bandlimited. We put $g = \mathcal{R}f$. Finally, we assume that the data are obtained scanning with *parallel standard geometry*, i.e. with a certain number p of equidistributed directions and taking in each of them $2q+1$ equispaced samples. The data have thus the following form:

$$\begin{aligned} &g_{j,l} := g(\theta_j, s_l) := \mathcal{R}f(\theta_j, s_l); \\ &\theta_j = e^{i\alpha_j}, \qquad \alpha_j = \frac{j\pi}{p}, \qquad j = 0,\dots,p-1, \qquad \Delta\alpha = \frac{\pi}{p}; \\ &s_l = \frac{l}{q}, \qquad l = -q,\dots,q, \qquad \Delta s = \frac{1}{q}. \end{aligned} \tag{29}$$

5.1 Filtered backprojection

Let $v \in \mathcal{S}(Z)$ and set $V = \mathcal{R}^\# v$. The convolutional identity in Prop. 3.18 was stated for Schwartz functions, but in our current setting, f belongs to $L^1(\mathbb{R}^n)$ (it is compactly supported), and is integrable over every line, the same computation extends by Fubini. Hence the following identity is valid pointwise for all $x \in \mathbb{R}^n$

$$\begin{aligned} (V*f)(x) = \big((\mathcal{R}^\# v) * f\big)(x) &\overset{\text{Prop. 3.18}}{=} \mathcal{R}^\#(v *_s g)(x) = \int_{S^{n-1}} (v *_s g)(\theta, x\cdot\theta)\, d\theta \\ &= \int_{S^{n-1}} \left(\int_{\mathbb{R}} v(\theta, x\cdot\theta - t)\, g(\theta,t)\, dt \right) d\theta. \end{aligned} \tag{30}$$

Notice that $V*f$ is bounded; hence $(1+|\cdot|)^{-N}(V*f) \in L^1(\mathbb{R}^n)$, $\forall N > n$. Thus $V*f$ defines a tempered distribution, which we identify with $V*f$ when needed.

The key feature of the filtered backprojection algorithm is the choice of a *point-spread function* V which plays the role of δ_0. Then $V * f$ approximates f, in a sense to be specified in each case. Once the corresponding filter v has been determined from $\mathcal{R}^{\#}v = V$, the right-hand side of (30) can be discretized.

Identity (30) explains the name of the algorithm: first the data g are filtered with v (this gives $v * g$) and then the backprojection $\mathcal{R}^{\#}$ is applied.

On both sides of (30), we have a convolution, which is computationally demanding. However, if we assume $V \in L^1(\mathbb{R}^n)$ then $\widehat{V}$ is well-defined in the usual Fourier sense, and then we'd have that $\widehat{V * f}(\zeta) = \widehat{V}(\zeta)\widehat{f}(\zeta)$ for $\zeta \in \mathbb{R}^n$: filtering could be performed by multiplication by $\widehat{V}$ in the frequency domain. The bad news is that $\widehat{V}$ might not be defined in the usual sense: even when $v \in \mathcal{S}(Z)$, the backprojection $V = \mathcal{R}^{\#}v$, need not belong to $L^1(\mathbb{R}^n)$. So, We will have to work distributionally: V will now be regarded as the element of $\mathcal{S}'$, and $\widehat{V}$ will be defined by duality: $\langle V, \varphi\rangle := \int_{\mathbb{R}^n} V(x)\varphi(x)\,dx$, $\langle \widehat{V}, \psi\rangle := \langle V, \widehat{\psi}\rangle$, for $\varphi, \psi \in \mathcal{S}(\mathbb{R}^n)$. Recall that $\widehat{V * f} \overset{\mathcal{S}'}{=} \widehat{V}\,\widehat{f}$ is equivalent to

$$\left\langle \widehat{V * f}, \varphi \right\rangle = \left\langle \widehat{V}\,\widehat{f}, \varphi \right\rangle = \left\langle \widehat{V}, \widehat{f}\,\varphi \right\rangle, \qquad \forall \varphi \in \mathcal{S}(\mathbb{R}^n),$$

since $\hat{f}$ is smooth and all derivatives of $\hat{f}$ have polynomial growth.[34] This way, it is valid to filter in the frequency domain.

Usually, V is chosen so that $V * f$ deletes or de-emphasizes high frequencies, which are mostly observation noise. Since f has essential bandwidth Ω, one looks for V such that

$$\widehat{V * f}(\zeta) \simeq \begin{cases} \widehat{f}(\zeta), & \text{if } |\zeta| \le \Omega, \\ 0, & \text{if } |\zeta| > \Omega. \end{cases}$$

In practice only radially symmetric functions $V(x) = V(|x|)$ are considered. Since $V(x) = \int_{S^{n-1}} v(\theta, x \cdot \theta)\,d\theta$, restricting our potential $v = v_Z \in \mathcal{S}(Z)$ to be independent of θ and even in s, so $v_Z(\theta, s) = v_0(s) = v_0(-s)$ for some even map $v_0 : \mathbb{R} \to \mathbb{R}$, gives a radial V. Also, we set $\widehat{v_Z}(\theta, \sigma) = \int_{\mathbb{R}} v_0(s)e^{-2\pi i\sigma s}\,ds = \widehat{v}_0(\sigma)$.

We now recall the following distributional identity [Nat01, Thm. II.1.4, p. 15]

$$\widehat{\mathcal{R}^{\#}h}(\zeta) = |\zeta|^{1-n}\left(\widehat{h}\left(\frac{\zeta}{|\zeta|}, |\zeta|\right) + \widehat{h}\left(-\frac{\zeta}{|\zeta|}, -|\zeta|\right)\right), \quad \zeta \neq 0,\ h \in \mathcal{S}(Z).^{35} \tag{31}$$

With our restrictions on v_Z we get $\widehat{v}_Z\left(\frac{\zeta}{|\zeta|}, |\zeta|\right) + \widehat{v}_Z\left(-\frac{\zeta}{|\zeta|}, -|\zeta|\right) = 2\widehat{v}_0(|\zeta|)$, since v_0 even implies $\widehat{v}_0$ even. And since $\mathcal{R}^{\#}v_Z = V$ (31) becomes:

$$\mathcal{F}_{\mathbb{R}^n} V(\zeta) = \widehat{V}(\zeta) = 2\,|\zeta|^{1-n}\,\widehat{v}_0\,(|\zeta|) = 2\,|\zeta|^{1-n}\,\mathcal{F}_{\mathbb{R}} v_0(|\zeta|), \qquad \zeta \neq 0.$$

[34] $\widehat{f} \in O_M(\mathbb{R}^2)$ because $f \in L^2(\mathbb{D})$ and is extended by zero outside $\mathbb{D}$, we have $x^\alpha f \in L^1(\mathbb{R}^2)$ for every multi-index α. Thus $\partial^\alpha \widehat{f}(\xi) = \int_{\mathbb{R}^2} (-2\pi i x)^\alpha f(x) e^{-2\pi i x\cdot\xi}\,dx$ is bounded for every α.

[35] Although the formula is written for $\zeta \neq 0$, the right-hand side is locally integrable and is therefore understood as defining a tempered distribution.

This equation suggests to define v, formally, such that

$$\widehat{v}(|\zeta|) := \frac{1}{2}\,|\zeta|^{n-1}\,\widehat{V}(\zeta), \qquad \zeta \in \mathbb{R}^n, \tag{32}$$

which makes sense since V radially symmetric implies $\widehat{V}$ is radially symmetric.

Now that we suspect how to obtain v from V, it is time to choose V. A particularly simple choice would be the ideal low-pass response $\widehat{V}_\Omega(\zeta) = \chi_{B(0,\Omega)}(\zeta)$. More generally, let $\widehat{\varphi} : \mathbb{R} \to \mathbb{C}$ be an even filter, close to 1 on $[-1,1]$ and vanishing outside $[-1,1]$. The previous discussion gives, for $\sigma = |\zeta|$

$$\widehat{V}_\Omega(\zeta) := \widehat{\varphi}\left(\frac{|\zeta|}{\Omega}\right), \quad \zeta \in \mathbb{R}^n \overset{(32)}{\Longrightarrow} \widehat{v}_\Omega(\sigma) = \frac{1}{2}|\sigma|^{n-1}\widehat{\varphi}\left(\frac{|\sigma|}{\Omega}\right), \quad \sigma \in \mathbb{R},$$

albeit in the case $n = 2$, v_Ω is not a Schwartz function.

Multiplication by $|\sigma|^{n-1}$ in the Fourier domain corresponds, up to a constant, to the operation ∂_s^{n-1} when n is odd, and to $H\partial_s^{n-1}$ when n is even. Thus (32) can be viewed as the Fourier counterpart of the Radon reconstruction formula (21).

In applications, different choices of $\widehat{\varphi}$ lead to different filters. The ideal choice $\widehat{\varphi} = \chi_{[-1,1]}$ gives a sharp cut-off, but sharp cut-offs tend to create oscillations. Smoother choices usually lose some resolution but reduce these artifacts. For this reason, the choice of the filter is partly theoretical and partly empirical [CMOCP07, Sec. 4.1]. In dimension $n = 2$, choices of $\widehat{\varphi}$ give the Ram–Lak, cosine and Shepp–Logan filters, shown in this order (we do not show the calculations for v_Ω):

$\widehat{\varphi}(\sigma)$	$v_\Omega(s)$	
$\chi_{[0,1]}(\sigma)$	$\Omega^2 u(2\pi\Omega s)$	$u(t) = \operatorname{sinc}(t) - \frac{1}{2}\operatorname{sinc}^2(t/2)$
$\cos\left(\frac{\pi\sigma}{2}\right)\chi_{[0,1]}(\sigma)$	$\frac{\Omega^2}{2}\left[u(2\pi\Omega s + \frac{\pi}{2}) + u(2\pi\Omega s - \frac{\pi}{2})\right]$	
$\operatorname{sinc}\left(\frac{\pi\sigma}{2}\right)\chi_{[0,1]}(\sigma)$	$\frac{2\Omega^2}{\pi}u(2\pi\Omega s)$,	$u(t) = \begin{cases} \frac{\pi/2 - t\sin t}{(\pi/2)^2 - t^2}, & t \neq \pm\pi/2, \\ \frac{1}{\pi}, & t = \pm\pi/2. \end{cases}$

We now discretize the identity (30). The continuous formula for $n = 2$ is

$$(V * f)(x) = \mathcal{R}^\#(v_\Omega *_s g)(x) = \int_0^{2\pi} (v_\Omega *_s g)(\theta, x\cdot\theta)\, d\alpha. \tag{33}$$

The first step is to discretize the convolution. Since $\operatorname{supp} f \subset B(0,1)$, for each fixed $\theta \in S^1$ we have $g(\theta,t) = 0$ for $|t| > 1$. Indeed, if $|t| > 1$, the line $\theta_t = \{x \in \mathbb{R}^2 : x\cdot\theta = t\}$ does not meet $B(0,1)$. Therefore

$$(v_\Omega *_s g)(\theta, s) = \int_{\mathbb{R}} v_\Omega(s-t)g(\theta,t)\, dt = \int_{-1}^{1} v_\Omega(s-t)g(\theta,t)\, dt.$$

According to the construction of the filter, v_Ω has bandwidth Ω. On the other hand, by Thm. 3.6, $\widehat{g}(\theta,\sigma) = \widehat{\mathcal{R}f}(\theta,\sigma) = \widehat{f}(\sigma\theta)$. Thus, if f is essentially Ω-bandlimited, then $g(\theta,\cdot)$ is essentially Ω-bandlimited as a function of s.

We would like to use Shannon's theorem, but it requires exact bandlimitation. We consider the frequency projection P_Ω in the variable s and rewrite g

$$\widehat{P_\Omega g}(\theta,\sigma) := \chi_{[-\Omega,\Omega]}(\sigma)\widehat{g}(\theta,\sigma) \implies g = P_\Omega g + (I - P_\Omega)g.$$

The part $P_\Omega g$ is exactly Ω-bandlimited, whereas $(I - P_\Omega)g$ contains the frequency tail. Substituting this decomposition into the convolution, we get

$$(v_\Omega * g)(\theta,s) = \int_{\mathbb{R}} v_\Omega(s-t)(P_\Omega g)(\theta,t)\,dt + \int_{\mathbb{R}} v_\Omega(s-t)((I-P_\Omega)g)(\theta,t)\,dt. \tag{34}$$

The second term is controlled using Cauchy–Schwarz, Plancherel and the essential bandlimitation of g:

$$\begin{aligned}\left|\int_{\mathbb{R}} v_\Omega(s-t)((I-P_\Omega)g)(\theta,t)\,dt\right| &\le \|v_\Omega(s-\cdot)\|_{L^2(\mathbb{R})}\|(I-P_\Omega)g(\theta,\cdot)\|_{L^2(\mathbb{R})} \\ &= \|v_\Omega\|_{L^2(\mathbb{R})}\left(\int_{|\sigma|>\Omega}|\widehat{g}(\theta,\sigma)|^2\,d\sigma\right)^{1/2}.\end{aligned} \tag{35}$$

Since $\operatorname{supp}(f) \subset B(0,1)$, for each fixed θ we have $\operatorname{supp} g(\theta,\cdot) \subset [-1,1]$. Thus, in order to use Shannon's Theorem on $P_\Omega g$, we take $\Delta s \le 1/2\Omega$ in (29) and neglect the frequency tail of g

$$\begin{aligned}(v_\Omega * g)(\theta,s) &\overset{(34)+(35)}{\simeq} \int_{\mathbb{R}} v_\Omega(s-t)(P_\Omega g)(\theta,t)\,dt \overset{\text{Shannon}}{=} \frac{1}{q}\sum_{l\in\mathbb{Z}} v_\Omega(s-s_l)\,(P_\Omega g)(\theta,s_l) \\ &\underset{\substack{\uparrow\\ \text{Discrete}\\ \text{analogue to (35)}}}{\simeq} \frac{1}{q}\sum_{l\in\mathbb{Z}} v_\Omega(s-s_l)\,g(\theta,s_l) \underset{\substack{\uparrow\\ g(\theta,s_l)=0\\ \text{for } |l|>q}}{=} \frac{1}{q}\sum_{l=-q}^{q} v_\Omega(s-s_l)\,g(\theta,s_l).\end{aligned} \tag{36}$$

The next step is to discretize the backprojection in (33). For fixed $x \in \mathbb{R}^2$, define

$$h_x(\alpha) := (v_\Omega *_s g)(e^{i\alpha}, x\cdot e^{i\alpha}).$$

This function is π-periodic[36]. A bandwidth estimate shows that h_x has essential bandwidth $4\pi\Omega$, in the sense that its Fourier coefficients

$$\widehat{h}_x(k) = \frac{1}{2\pi}\int_0^{2\pi}(v_\Omega *_s g)(e^{i\alpha}, x\cdot e^{i\alpha})e^{-ik\alpha}\,d\alpha$$

[36] $h_x(\alpha+\pi) = (v_\Omega *_s g)(e^{i(\alpha+\pi)}, x\cdot e^{i(\alpha+\pi)}) = (v_\Omega *_s g)(-\theta,-x\cdot\theta) = \int_{\mathbb{R}} v_\Omega(-x\cdot\theta - t)\,g(-\theta,t)\,dt$
$\overset{u=-t}{=} \int_{\mathbb{R}} v_\Omega(-x\cdot\theta+u)\,g(-\theta,-u)\,du = \int_{\mathbb{R}} v_\Omega(x\cdot\theta-u)\,g(\theta,u)\,du = (v_\Omega *_s g)(\theta, x\cdot\theta).$

are negligible for $|k| > 4\pi\Omega$ [CMOCP07, Sec. 4.1, pp. 27–28]. Using the same approach as before, again up to another negligible error, we may use the angular sampling rule. Taking $\Delta\alpha = 1/(2\Omega)$ in (29),

$$\theta_j = e^{ij\pi/p}, \qquad j = 0, \dots, 2p-1, \qquad \Delta\alpha = \frac{\pi}{p} \le \frac{1}{2\Omega},$$

and using $2p$ nodes in order to cover the full interval $[0, 2\pi)$ with this step, we use the $2p$ nodes $\theta_j = e^{ij\pi/p}$, $j = 0, \dots, 2p-1$ ($j = 2p$ is not included since it gives the same direction as $j = 0$), we have

$$\begin{aligned}(V * f)(x) &= \int_0^{2\pi} (v_\Omega *_s g)(\theta, x\cdot\theta)\, d\alpha \simeq \frac{\pi}{p} \sum_{j=0}^{2p-1} (v_\Omega *_s g)(\theta_j, x\cdot\theta_j) \\ &= \frac{2\pi}{p} \sum_{j=0}^{p-1} (v_\Omega *_s g)(\theta_j, x\cdot\theta_j),\end{aligned} \tag{37}$$

where the last identity follows from π-periodicity.

Combining (36) and (37), and always taking $\max\{\Delta\alpha, \Delta s\} \le \frac{1}{2\Omega}$, we obtain the direct discrete filtered-backprojection formula

$$(V * f)(x) \simeq \frac{2\pi}{p} \sum_{j=0}^{p-1} \Delta s \sum_{l=-q}^{q} v_\Omega(x\cdot\theta_j - s_l) g(\theta_j, s_l). \tag{38}$$

This is the discrete version of the continuous identity (33): for each direction θ_j, one filters the measured data $g(\theta_j, \cdot)$ in the s-variable and then backprojects the result by evaluating it at $s = x\cdot\theta_j$.

5.1.1 Cost. Formula (38) is conceptually simple but computationally expensive. For each fixed point x its cost is $O(pq)$, and using that $p \ge 2\pi\Omega$ and $q \ge 2\Omega$, we get $O(\Omega^2)$ per point. Since f has essential bandwidth Ω, one must reconstruct it on a spatial lattice with stepsize $\frac{1}{2\Omega}$. In dimension 2, this gives $O(\Omega^2)$ reconstruction points. Therefore the total cost of the direct formula is $O(\Omega^4)$.

To improve this cost, for each fixed direction θ_j, set $H_j(t) := (v_\Omega *_s g)(\theta_j, t)$. For $j = 0, \dots, p-1$ and $k = -q, \dots, q$, we have

$$h_{j,k} := H_j(s_k) = (v_\Omega *_s g)(\theta_j, s_k) \simeq \Delta s \sum_{l=-q}^{q} v_\Omega(s_k - s_l) g_{j,l}.$$

For fixed j and k, this sum costs $O(q)$ operations. For each fixed j, there are $O(q)$ values of k, thus computing all the values $h_{j,k}$ for that direction θ_j costs $O(q^2)$. Repeating this for the p directions gives a precomputation cost $O(pq^2) \simeq O(\Omega^3)$. Once these values have been computed, $H_j(x\cdot\theta_j)$ is no longer computed by a new convolution sum, but by linear interpolation. For $t = x\cdot\theta_j$, set

$$\alpha = \alpha(j,x) := \lfloor qt \rfloor, \qquad \beta = \beta(j,x) := qt - \lfloor qt \rfloor.$$

Then $0 \le \beta < 1$ and $\frac{\alpha}{q} \le t < \frac{\alpha+1}{q}$. Thus $H_j(x \cdot \theta_j) \simeq (1-\beta)h_{j,\alpha} + \beta h_{j,\alpha+1}$. At the endpoint $qt = q$, we take $\alpha = q-1$ and $\beta = 1$. Therefore the final approximation is

$$f_A(x) = \frac{2\pi}{p} \sum_{j=0}^{p-1} \left((1-\beta(j,x))h_{j,\alpha(j,x)} + \beta(j,x)h_{j,\alpha(j,x)+1} \right).$$

The backprojection step now costs $O(p)$ operations for each reconstructed point. Since there are $O(\Omega^2)$ reconstruction points, this step costs $O(\Omega^3)$.

Together with the precomputation cost $O(pq^2) \simeq O(\Omega^3)$, the whole method has complexity $O(\Omega^3)$, instead of the $O(\Omega^4)$ cost of the direct formula.

5.1.2 Algorithm. *Step 1.* For every direction θ_j, compute the discrete convolution

$$h_{j,k} = \frac{1}{q} \sum_{l=-q}^{q} v_\Omega \left(\frac{k-l}{q} \right) g_{j,l}, \qquad k = -q, \ldots, q.$$

Step 2 For each reconstruction point x, compute the discrete backprojection by linear interpolation. Namely, for each j, set

$$\alpha = \alpha(j,x) := \lfloor q\, x \cdot \theta_j \rfloor, \qquad \beta = \beta(j,x) := q\, x \cdot \theta_j - \lfloor q\, x \cdot \theta_j \rfloor.$$

Then

$$f_A(x) := \frac{2\pi}{p} \sum_{j=0}^{p-1} \left((1-\beta(j,x))h_{j,\alpha(j,x)} + \beta(j,x)h_{j,\alpha(j,x)+1} \right).$$

At the endpoint $q\, x \cdot \theta_j = q$, we take $\alpha(j,x) = q-1$ and $\beta(j,x) = 1$.

5.2 The gridding method

The steps of the method can be informally summarized in the following chain:

$$g(\theta_j, s_l) \rightsquigarrow \widehat{g}(\theta_j, \sigma_l) = \widehat{f}(\sigma_l \theta_j) \rightsquigarrow \widehat{W} * \widehat{f} = \widehat{Wf} \rightsquigarrow z_k \simeq \widehat{Wf}\left(\frac{k}{2}\right) \rightsquigarrow (Wf)\left(\frac{m}{q}\right) \rightsquigarrow f\left(\frac{m}{q}\right),$$

where $W : \mathbb{R}^2 \to \mathbb{R}$ is a so-called *weight* function and z_k is the estimation of the evaluation of $\widehat{Wf}$ on a 2D Cartesian grid.

Usually W is close to 1 in the reconstruction region $\mathbb{D} = B(0,1)$, vanishing when $|x| \ge a > 1$, and such that $\widehat{W}$ is concentrated near 0. Then Wf is close to f in the reconstruction region. Therefore, reconstructing f is reduced to approximating $\widehat{Wf}$. For $\zeta \in \mathbb{R}^2$, we have

$$\begin{aligned} \widehat{Wf}(\zeta) = (\widehat{W} * \widehat{f})(\zeta) &= \int_{\mathbb{R}^2} \widehat{W}(\zeta - \eta)\widehat{f}(\eta)\, d\eta = \int_0^\infty \int_{S^1} \widehat{W}(\zeta - \sigma\theta)\widehat{f}(\sigma\theta)\sigma\, d\theta d\sigma \\ &= \int_0^\infty \int_{S^1} \widehat{W}(\zeta - \sigma\theta)\widehat{g}(\theta,\sigma)\sigma\, d\theta d\sigma. \end{aligned} \tag{39}$$

Since Wf is supported in $\mathbb{D}$, which is symmetric, the function $\widehat{Wf}$ has bandwidth 1. By Shannon's theorem, $\widehat{Wf}$ is thus determined by its values on the Cartesian lattice $(1/2)\mathbb{Z}^2$. Once these values have been approximated, we apply inverse DFT, to obtain

$$(Wf)(m/q) \simeq \frac{1}{4} \sum_{|k|\le q} \widehat{Wf}(k/2)\, e^{2\pi i \frac{k}{2}\cdot\frac{m}{q}}.$$

Finally, f is recovered by dividing pointwise by W.

The success of the method lies in a suitable discretization of (39). The values θ_j used in the discretization are determined by the measured directions, since the FT of g is taken only with respect to the second variable. Thus, using the trapezoidal rule[37] with stepsizes $\Delta\varphi = \pi/p$ in the angular variable and $\Delta\sigma$ in the radial variable, we obtain the first approximation of $\widehat{Wf}\left(\frac{k}{2}\right)$ as in (39):

$$\begin{aligned}\int_0^\infty \int_{S^1} \widehat{W}(\zeta - \sigma\theta)\widehat{g}(\theta,\sigma)\sigma\, d\theta d\sigma &\simeq \int_0^\infty \left(\frac{\pi}{p}\sum_{j=0}^{2p-1} \widehat{W}\left(\frac{k}{2} - \sigma\theta_j\right)\widehat{g}(\theta_j,\sigma)\right)\sigma\, d\sigma \\ &\simeq \Delta\sigma \sum_{l=0}^{\infty}\left(\frac{\pi}{p}\sum_{j=0}^{2p-1} \widehat{W}\left(\frac{k}{2} - l\Delta\sigma\,\theta_j\right)\widehat{g}(\theta_j, l\Delta\sigma)\right)(l\Delta\sigma)\end{aligned}$$

obtaining

$$\widehat{Wf}\left(\frac{k}{2}\right) \simeq \sum_{l=0}^{\infty}\sum_{j=0}^{2p-1} \widehat{W}\left(\frac{k}{2} - l\Delta\sigma\,\theta_j\right)\widehat{g}(\theta_j, l\Delta\sigma)\,(l\Delta\sigma)\,\frac{\pi}{p}\,\Delta\sigma. \tag{40}$$

In the derivation of (40), however, we have ignored that the radial integral is defined only on $[0,\infty)$, not on the whole real line. This produces the annihilation of all the information stored in the cell represented by $l = 0$, namely the information given by $\widehat{g}(\theta_j, 0)$. This produces artifacts and must be corrected.

5.2.1 Discretization of bandlimited functions on $[0,\infty)$ The Poisson summation formula for the rectangular lattice $\Delta\sigma\,\mathbb{Z}$ [Grö01, Prop. 1.4.2 and Eq. (1.37)] $\sum_l \varphi(l) = \frac{1}{\Delta\sigma}\sum_l \widehat{\varphi}\left(\frac{l}{\Delta\sigma}\right)$ for $\varphi \in \mathcal{S}(\mathbb{R})$ is equivalent, by $\breve{\varphi}(s) = \widehat{\varphi}(-s)$, to

$$\sum_{l\in\mathbb{Z}} \varphi(l\Delta\sigma) \overset{\text{Poisson summation}}{\downarrow} \frac{1}{\Delta\sigma}\sum_{l\in\mathbb{Z}} \breve{\varphi}\left(\frac{l}{\Delta\sigma}\right) \overset{*}{=} \frac{1}{\Delta\sigma}\int_{\mathbb{R}} \varphi(\sigma)\, d\sigma + \frac{1}{\Delta\sigma}\sum_{l\in\mathbb{Z}\setminus\{0\}} \breve{\varphi}\left(\frac{l}{\Delta\sigma}\right),$$

[37] If $t_l = a + l\Delta t$, $l = 0,\dots,N$ and $\Delta t = \frac{b-a}{N}$, then

$$\int_a^b H(t)\, dt \simeq \Delta t\left(\frac{1}{2}H(t_0) + \sum_{l=1}^{N-1} H(t_l) + \frac{1}{2}H(t_N)\right).$$

For periodic and half-line integrals we shall use the approximations

$$\int_{S^1} H(\theta)\, d\theta \simeq \frac{\pi}{p}\sum_{j=0}^{2p-1} H(e^{ij\pi/p}), \qquad \int_0^\infty B(\sigma)\, d\sigma \simeq \Delta\sigma \sum_{l\ge 0} B(l\Delta\sigma).$$

where in $*$ we used $\check{\varphi}(0) = \int_{\mathbb{R}} \varphi(\sigma) e^{2\pi i \sigma \cdot 0} d\sigma$. Hence

$$\int_{\mathbb{R}} \varphi(\sigma)\, d\sigma = \Delta\sigma \sum_{l\in\mathbb{Z}} \varphi(l\Delta\sigma) - \sum_{l\in\mathbb{Z}\setminus\{0\}} \check{\varphi}\left(\frac{l}{\Delta\sigma}\right). \tag{41}$$

Let u be an auxiliary function, and set $\varphi(\sigma) := u(\sigma)\,\mathrm{sgn}(\sigma)$, with $u(0) = 0$. Although u sgn need not belong to $\mathcal{S}(\mathbb{R})$, the following computation is understood in the usual regularized sense. Recalling Hilbert's operator, we obtain

$$\widehat{\varphi} = \varphi(\zeta) = \mathrm{sgn}(\zeta) u(\zeta) = \mathrm{sgn}(\zeta)\widehat{u}\,(\zeta) \overset{\text{Eq. (20)}}{=} \widehat{H\check{u}}\,(\zeta) \tag{42}$$

Thus $\check{\varphi} = H\check{u}$, and applying (formally) (41) to φ gives

$$\int_{\mathbb{R}} u(\sigma)\,\mathrm{sgn}(\sigma)\, d\sigma = \Delta\sigma \sum_{l\in\mathbb{Z}} u(l\Delta\sigma)\,\mathrm{sgn}(l\Delta\sigma) - r, \underbrace{\sum_{l\in\mathbb{Z}\setminus\{0\}} (H\check{u})\left(\frac{l}{\Delta\sigma}\right)}_{r}.$$

Using the integral formula $(H\psi)(t) = \frac{i}{\pi}\,\mathrm{pv}\int_{\mathbb{R}} \psi(t-s)/s\, ds$, the Fourier identity $\check{u}(s)\, s = -\frac{1}{2\pi i}\mathcal{F}^{-1}(u')(s)$, and grouping the terms l and $-l$, we obtain

$$\begin{aligned} r &= \frac{i}{\pi} \sum_{l\in\mathbb{Z}\setminus\{0\}} \int_{\mathbb{R}} \frac{\check{u}(\frac{l}{\Delta\sigma} - s)}{s}\, ds \overset{t=\frac{l}{\Delta\sigma}-s}{=} \frac{i}{\pi} \sum_{l\in\mathbb{Z}\setminus\{0\}} \int_{\mathbb{R}} \frac{\check{u}(t)}{\frac{l}{\Delta\sigma} - t}\, dt \\ &= \frac{i}{\pi} \sum_{l=1}^{\infty} \int_{\mathbb{R}} \check{u}(t) \left(\frac{1}{\frac{l}{\Delta\sigma} - t} + \frac{1}{-\frac{l}{\Delta\sigma} - t} \right) dt = \frac{i}{\pi} \sum_{l=1}^{\infty} \int_{\mathbb{R}} \frac{2t\,\check{u}(t)}{\left(\frac{l}{\Delta\sigma}\right)^2 - t^2}\, dt \\ &= \frac{i}{\pi} \sum_{l=1}^{\infty} \int_{\mathbb{R}} \frac{2\left(-\frac{1}{2\pi i}\check{u}'(t)\right)}{\left(\frac{l}{\Delta\sigma}\right)^2 - t^2}\, dt = -\frac{1}{\pi^2} \sum_{l=1}^{\infty} \int_{\mathbb{R}} \frac{\check{u}'(t)}{\left(\frac{l}{\Delta\sigma}\right)^2 - t^2}\, dt, \end{aligned} \tag{43}$$

where the integral is understood as a principal value integral.

Now suppose u has bandwidth A, then $\mathcal{F}^{-1}(u')$ is supported in $[-A, A]$. Assume also that $\Delta\sigma \ll \frac{1}{A}$. Then, for $t \in [-A, A]$, the term t^2 in the denominator of (43) is negligible with respect to $(l/\Delta\sigma)^2$; i.e. if $a := \left(\frac{t\Delta\sigma}{l}\right)^2 \ll 1$ then $r = -\frac{1}{6}(\Delta\sigma)^2 u'(0)$.[38] Therefore if $u(0) = 0$,

$$\int_{\mathbb{R}} u(\sigma)\,\mathrm{sgn}(\sigma)\, d\sigma \simeq \Delta\sigma \sum_{l\in\mathbb{Z}} u(l\Delta\sigma)\,\mathrm{sgn}(l\Delta\sigma) + \frac{1}{6}(\Delta\sigma)^2 u'(0). \tag{44}$$

[38] We use that if $a = \left(\frac{t\Delta\sigma}{l}\right)^2 \ll 1$ then $(1-a)^{-1} \simeq 1$, obtaining

$$r = -\frac{1}{\pi^2} \sum_{l=1}^{\infty} \int_{-A}^{A} \check{u}'(t)\, \frac{1}{\left(\frac{l}{\Delta\sigma}\right)^2} \frac{1}{1 - \left(\frac{t\Delta\sigma}{l}\right)^2}\, dt \simeq -\frac{1}{\pi^2} \sum_{l=1}^{\infty} \frac{1}{\left(\frac{l}{\Delta\sigma}\right)^2} \int_{-A}^{A} \check{u}'(t)\, dt = -\frac{1}{\pi^2}(\Delta\sigma)^2 \sum_{l=1}^{\infty} \frac{1}{l^2} \int_{\mathbb{R}} \check{u}'(t) e^{2\pi i t \cdot 0}\, dt$$

$$\overset{\mathcal{F}^{-1}}{=} -\frac{1}{\pi^2}(\Delta\sigma)^2 \sum_{l=1}^{\infty} \frac{1}{l^2} u'(0) = -\frac{1}{\pi^2}(\Delta\sigma)^2 \frac{\pi^2}{6} u'(0) = -\frac{1}{6}(\Delta\sigma)^2 u'(0).$$

We apply this formula to $u(\sigma) := \sigma G(\sigma)$, with $G(\sigma) := \int_{S^1} \widehat{W}(\zeta - \sigma\theta)\widehat{g}(\theta,\sigma)\, d\theta$ is the even function in (39). Since G is even and $u'(0) = G(0)$, formula (44) gives

$$\int_0^\infty \sigma G(\sigma)\, d\sigma \simeq \Delta\sigma \sum_{l=1}^{\infty} l\Delta\sigma\, G(l\Delta\sigma) + \frac{(\Delta\sigma)^2}{12} G(0) = \Delta\sigma \sum_{l=0}^{\infty} \sigma_l G(l\Delta\sigma),$$

where $\sigma_l = \begin{cases} l\Delta\sigma, & l>0, \\ \Delta\sigma/12, & l=0. \end{cases}$ In our case, this means that

$$\int_0^\infty \int_{S^1} \widehat{W}(\zeta - \sigma\theta)\widehat{g}(\theta,\sigma)\sigma\, d\theta\, d\sigma \simeq \Delta\sigma \sum_{l=0}^{\infty} \sigma_l \int_{S^1} \widehat{W}(\zeta - l\Delta\sigma\,\theta)\widehat{g}(\theta, l\Delta\sigma)\, d\theta.$$

Thus we obtain a corrected discretization of $\widehat{Wf}(k/2)$:

$$\begin{aligned} \widehat{Wf}\left(\frac{k}{2}\right) &\simeq \Delta\sigma \sum_{l=0}^{\infty} \sigma_l \int_{S^1} \widehat{W}\left(\frac{k}{2} - l\Delta\sigma\,\theta\right) \widehat{g}(\theta, l\Delta\sigma)\, d\theta \\ &\simeq \Delta\sigma \sum_{l=0}^{\infty} \sigma_l \sum_{j=0}^{2p-1} \widehat{W}\left(\frac{k}{2} - l\Delta\sigma\,\theta_j\right) \widehat{g}(\theta_j, l\Delta\sigma)\frac{\pi}{p}. \end{aligned}$$

This discretization differs from (40) only in the term $l = 0$. Precisely that term corrects the Riemann sum for the error introduced by truncating the integration domain at 0.

We now choose the radial step $\Delta\sigma$. By (41), the error of a discretization of a function φ is controlled by the values $\check{\varphi}\left(l/(\Delta\sigma)\right)$, $l \neq 0$. If φ has bandwidth B, then supp $\widehat{\varphi} \subset [-B, B]$. Since $\check{\varphi}(t) = \widehat{\varphi}(-t)$, this also gives supp $\check{\varphi} \subset [-B, B]$. Thus, if $1/\Delta\sigma > B$, then $\check{\varphi}\left(l/(\Delta\sigma)\right) = 0$, $l \neq 0$.

In the present case we estimate the bandwidth, as a function of σ, of $\sigma\,\psi(\sigma)$ with $\psi(\sigma) := \widehat{W}(\zeta - \sigma\theta)\widehat{g}(\theta,\sigma)$. Since $\mathcal{F}_\sigma^{-1}\left(\sigma\psi(\sigma)\right) = \frac{1}{2\pi i}\frac{d}{ds}\check{\psi}(s)$, we have that the factor σ does not enlarge the bandwidth. Now

$$\check{\psi} = \mathcal{F}_\sigma^{-1}\Big(\widehat{W}(\zeta - \sigma\theta)\widehat{g}(\theta,\sigma)\Big) = \mathcal{F}_\sigma^{-1}\Big(\widehat{W}(\zeta - \sigma\theta)\Big) * g(\theta,\cdot).$$

Recall that $\mathrm{supp}(A * B) \subset \overline{\mathrm{supp}(A) + \mathrm{supp}(B)}$.[39] In our case, $\sigma \mapsto \widehat{g}(\theta,\sigma)$ has bandwidth 1, since $g(\theta,\cdot)$ is supported in $[-1,1]$ (because supp $f \subset B(0,1)$).

It remains to estimate the bandwidth of $\sigma \longmapsto \widehat{W}(\zeta - \sigma\theta)$, i.e. the support of $\mathcal{F}_\sigma^{-1}\Big(\widehat{W}(\zeta - \sigma\theta)\Big)$. We will show it is $[-a, a]$ (recall that we had set $W(x) = 0$ for $|x| > a$). We use the distributional identity $\delta(x\cdot\theta + s) = \int_{\mathbb{R}} e^{2\pi i(x\cdot\theta+s)\sigma}\, d\sigma$. Indeed,

[39] Let $x_0 \notin \overline{\mathrm{supp}\, A + \mathrm{supp}\, B}$. Then there exists a neighborhood U of x_0 such that $U \cap (\mathrm{supp}\, A + \mathrm{supp}\, B) = \varnothing$. Hence, for every $x \in U$ and every y, $A(y)B(x-y) = 0$, since otherwise $y \in \mathrm{supp}\, A$ and $x - y \in \mathrm{supp}\, B$, so $x \in \mathrm{supp}\, A + \mathrm{supp}\, B$. Thus $(A * B)(x) = \int A(y)B(x-y)\, dy = 0$ for $x \in U$. Therefore $x_0 \notin \mathrm{supp}(A * B)$.

thus for $|s| > a$, we calculate $\mathcal{F}_\sigma^{-1}\left(\widehat{W}(\zeta - \sigma\theta)\right)(s)$:

$$\begin{aligned}\mathcal{F}_\sigma^{-1}\left(\widehat{W}(\zeta - \sigma\theta)\right)(s) &= \int_{\mathbb{R}} \widehat{W}(\zeta - \sigma\theta) e^{2\pi i s\sigma}\, d\sigma \\ &= \int_{\mathbb{R}} \left(\int_{|x|\le a} W(x) e^{-2\pi i x\cdot(\zeta-\sigma\theta)}\, dx \right) e^{2\pi i s\sigma}\, d\sigma \\ &= \int_{|x|\le a} W(x) e^{-2\pi i x\cdot\zeta} \left(\int_{\mathbb{R}} e^{2\pi i (x\cdot\theta + s)\sigma}\, d\sigma \right) dx \\ &= \int_{|x|\le a} W(x) e^{-2\pi i x\cdot\zeta} \delta(x\cdot\theta + s)\, dx = 0.\end{aligned}$$

The last equality follows because, if $|x| \le a$, then $|x\cdot\theta| \le a < |s|$, so the hyperplane $x\cdot\theta = -s$ does not meet supp W. Therefore ψ has bandwidth at most $a+1$. This suggests the choice $\Delta\sigma \ll \frac{1}{a+1}$.

5.2.2 Algorithm Once we have chosen W with the requirements shown above, we proceed as follows. We set $\Delta\sigma \ll 1/(a+1)$. Recall the variable names we set before in (29).

Step 1. For each measured direction θ_j, we approximate the one-dimensional FT of $g(\theta_j, \cdot)$ at the frequencies $\sigma_\nu = \nu\Delta\sigma$, with $|\nu|\Delta\sigma \le \Omega$, $\nu \in \mathbb{Z}$. Recall we had set $s_l = \frac{l}{q}$ and $\Delta s = \frac{1}{q}$. Thus,

$$\widehat{g}(\theta_j, \nu\Delta\sigma) = \int_{\mathbb{R}} g(\theta_j, s) e^{-2\pi i\, \nu\Delta\sigma\, s}\, ds \simeq \frac{1}{q} \sum_{l=-q}^{q} g(\theta_j, \tfrac{l}{q}) e^{-2\pi i\, \nu\Delta\sigma\, \frac{l}{q}}.$$

Thus, using $g_{j,l} = g(\theta_j, \frac{l}{q})$ from (29), we set

$$\widehat{g}_{j,\nu} := \frac{1}{q} \sum_{l=-q}^{q} g_{j,l} e^{-2\pi i\, \nu\Delta\sigma\, \frac{l}{q}} \simeq \widehat{g}(\theta_j, \nu\Delta\sigma), \qquad j = 0, \dots, p-1, \quad |\nu|\Delta\sigma \le \Omega.$$

For the opposite directions we use the Radon symmetry: $\theta_{j+p} := -\theta_j$, $\widehat{g}_{j+p,\nu} := \widehat{g}_{j,-\nu}$, $j = 0, \dots, p-1$. Indeed, $g(-\theta, s) = g(\theta, -s)$, and therefore $\widehat{g}(-\theta, \sigma) = \widehat{g}(\theta, -\sigma)$.

Step 2. For every $k \in \mathbb{Z}^2$, $|k| \le q$, compute the discretization of $\widehat{Wf}(k/2)$:

$$\widehat{Wf}\left(\frac{k}{2}\right) \simeq z_k := \frac{\pi}{p}\Delta\sigma \sum_{\substack{\nu\ge 0 \\ \nu\Delta\sigma\le\Omega}} \sum_{j=0}^{2p-1} \sigma_\nu \widehat{W}\left(k/2 - \nu\Delta\sigma\,\theta_j\right) \widehat{g}_{j,\nu}, \quad \sigma_\nu = \begin{cases} \nu\Delta\sigma, & \nu > 0, \\ \Delta\sigma/12, & \nu = 0. \end{cases}$$

Since $\widehat{W}$ is concentrated near 0 and decays rapidly, only few terms are relevant in this sum, namely those for which $\left|k/2 - \nu\Delta\sigma\,\theta_j\right| \le C$.

Step 3. Since $z_k \simeq \widehat{Wf}(k/2)$, the inverse discrete Fourier formula gives values of Wf on the spatial grid:

$$(Wf)\left(\frac{m}{q}\right) \simeq \sum_{|k|\le q} e^{2\pi i\, k\cdot \frac{m}{q}} z_k.$$

Finally, since W is nonzero in the reconstruction region, we divide pointwise by W and define

$$f_m := \frac{1}{W\left(\frac{m}{q}\right)} \sum_{|k|\le q} e^{2\pi i\, k\cdot \frac{m}{q}} z_k, \qquad m \in \mathbb{Z}^2, \quad |m| \le q.$$

5.2.3 Cost. Step 1 would cost $O(pq\,\Omega/\Delta\sigma)$ if computed directly. Since, for each fixed θ_j, it is a one-dimensional Fourier transform of $g_{j,l} = g(\theta_j, s_l)$ in the s-variable, it can be computed by FFT, with cost $O(pq\log q)$. Step 2 is local and costs $O(q^2)$. Finally, Step 3 is written as a two-dimensional inverse DFT on the Cartesian grid, and can therefore be evaluated by FFT, with cost $O(q^2\log q)$. Hence

$$C = O(pq\log q) + O(q^2) + O(q^2\log q) = O(\Omega^2\log\Omega) \qquad \text{when } p \simeq q \simeq \Omega.$$

6 Conclusion

This work has presented the Radon transform as the mathematical core of computed tomography, moving from its continuous theory to its discretization and numerical reconstruction. The Fourier slice theorem and the inversion formulas explain why projection data determine the original object, while the filtered backprojection and gridding methods show how this principle can be turned into practical algorithms. Thus, CT illustrates how abstract Fourier analysis becomes an effective tool for recovering images from indirect measurements.

Appendix A. The concentration operator

A.1 Additional functional analysis preliminaries. Let H be a Hilbert space with a countable orthonormal basis (e_k). A bounded operator $T : H \to H$ is called Hilbert–Schmidt if

$$\|T\|_{\mathrm{HS}}^2 := \sum_k \|Te_k\|^2 < \infty.$$

This condition is independent of the chosen orthonormal basis [Grö01, App. A.8, p. 331].

An operator $A : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is called an integral operator with kernel K if K is measurable and, for every $f \in L^2(\mathbb{R}^n)$, Af has a representative satisfying, for a.e. y,

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

If $K \in L^2(\mathbb{R}^n \times \mathbb{R}^n)$, then A is Hilbert–Schmidt and $\|A\|_{\mathrm{HS}} = \|K\|_{L^2(\mathbb{R}^n \times \mathbb{R}^n)}$ [Grö01, App. A.8, p. 331]. In particular, A is compact [RS80, Thm. VI.23, pp. 210–211].

Finally, if $T \geq 0$ is a bounded operator on a separable Hilbert space, then

$$\mathrm{tr}(T) := \sum_k \langle Te_k, e_k \rangle \in [0, \infty]$$

is independent of the chosen orthonormal basis (e_k) [RS80, Thm. VI.18, p. 207]. Hence, if T is positive compact with eigenvalues (λ_k), counted with multiplicity, then

$$\mathrm{tr}(T) = \sum_k \lambda_k.$$

A.2 Proof of Prop. 4.3. We prove this proposition in two separate parts. Recall the definition of the concentration operator.

$$C_{\Omega,R} : PW_\Omega \to PW_\Omega, \qquad C_{\Omega,R} f := P_\Omega(\chi_R \cdot f), \qquad P_\Omega f := \mathcal{F}^{-1}(\chi_\Omega \cdot \mathcal{F}(f)).$$

Proposition A.1. ***Basic spectral properties.*** *The operator $C_{\Omega,R}$ is linear, bounded, compact, self-adjoint, positive and injective. Then $\int_{B_R} |f(x)|^2\, dx = \langle C_{\Omega,R} f, f \rangle$ and if $\lambda_k(\Omega, R)$ for $k \in \mathbb{N}$ are its eigenvalues, listed in non-increasing order and counted with multiplicity, we have $\lambda_k(\Omega, R) \in (0,1)$ and:*

$$\mathrm{tr}(C_{\Omega,R}) = \sum_{k=1}^{\infty} \lambda_k = D_{\Omega,R},$$
$$\mathrm{tr}(C_{\Omega,R}^2) = \sum_{k=1}^{\infty} \lambda_k^2 = D_{\Omega,R} \pm o\big((R\Omega)^n\big), \qquad R\Omega \to \infty.$$

Proof. For simplicity, write $C_{\Omega,R}f := P_\Omega(\chi_R f)$, and $P_\Omega f := \mathcal{F}^{-1}\big(\chi_\Omega \mathcal{F}(f)\big)$ and $C = C_{\Omega,R}$. Observe that for the proof, we omit the dot for pointwise multiplication.

First, C is linear since it is the composition of linear operators.

Second, P_Ω is the orthogonal projection onto PW_Ω. Indeed, for $h \in L^2(\mathbb{R}^n)$,

$$\widehat{P_\Omega h} = \mathcal{F}\left[\mathcal{F}^{-1}\big(\chi_\Omega \mathcal{F}(h)\big)\right] = \chi_\Omega \widehat{h}.$$

Thus $P_\Omega h \in PW_\Omega$. Moreover, if $g \in PW_\Omega$, then

$$\langle h - P_\Omega h, g\rangle = \left\langle \widehat{h} - \chi_\Omega \widehat{h}, \widehat{g}\right\rangle = \int_{\mathbb{R}^n} (1 - \chi_\Omega(\xi))\widehat{h}(\xi)\overline{\widehat{g}(\xi)}\, d\xi = 0,$$

because $\widehat{g}$ is supported in B_Ω. Hence $h - P_\Omega h$ is orthogonal to PW_Ω. In particular, orthogonal projections have norm at most one, and therefore

$$\|Cf\|_2 = \|P_\Omega(\chi_R f)\|_2 \le \|\chi_R f\|_2 \le \|f\|_2 .$$

Thus C is bounded.

Third, for $f, g \in PW_\Omega$, using that P_Ω is self-adjoint and that multiplication by χ_R is self-adjoint, since χ_R is real-valued, we get

$$\begin{aligned}\langle Cf, g\rangle = \langle P_\Omega(\chi_R f), g\rangle = \langle \chi_R f, P_\Omega g\rangle &\overset{P_\Omega g = g}{=} \langle \chi_R f, g\rangle = \langle f, \chi_R g\rangle \\ &\overset{P_\Omega f = f}{=} \langle P_\Omega f, \chi_R g\rangle = \langle f, P_\Omega(\chi_R g)\rangle = \langle f, Cg\rangle .\end{aligned} \tag{45}$$

Hence C is self-adjoint.

Fourth, taking $g = f$ in the previous computation gives

$$\langle Cf, f\rangle = \langle \chi_R f, f\rangle = \int_{\mathbb{R}^n} \chi_R(x)|f(x)|^2\, dx = \int_{B_R} |f(x)|^2\, dx. \tag{46}$$

Therefore C is positive.

Fifth, if $C_{\Omega,R}f = 0$, then by (46), $\int_{B_R} |f(x)|^2\, dx = 0$. Thus $f = 0$ almost everywhere on B_R. Since $f \in PW_\Omega$, f has a continuous representative, and therefore $f = 0$ on B_R. But Paley–Wiener functions are real analytic; hence, by analyticity, $f \equiv 0$. Thus C is injective.

Sixth, to see compactness, write C as an integral operator. Notice that we can use Fubini's theorem, since

$$\begin{aligned}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \chi_\Omega(\xi)\, |\chi_R(x) f(x)|\, dx d\xi &= |B_\Omega|\, \|\chi_R f\|_1 \\ &\overset{\text{C.S.}}{\le} |B_\Omega||B_R|^{1/2}\, \|f\|_2 < \infty.\end{aligned}$$

Hence

$$\begin{aligned} Cf(y) &= \mathcal{F}^{-1}\big(\chi_\Omega \widehat{\chi_R f}\big)(y) = \int_{\mathbb{R}^n} \chi_\Omega(\xi) \left(\int_{\mathbb{R}^n} \chi_R(x) f(x) e^{-2\pi i x\cdot\xi}\, dx \right) e^{2\pi i y\cdot\xi}\, d\xi \\ &= \int_{\mathbb{R}^n} f(x)\chi_R(x) \left(\int_{\mathbb{R}^n} \chi_\Omega(\xi) e^{2\pi i (y-x)\cdot\xi}\, d\xi \right) dx \\ &= \int_{\mathbb{R}^n} f(x) \underbrace{\chi_R(x)\widetilde{\chi_\Omega}(y-x)}_{K_{\Omega,R}(x,y)}\, dx = \int_{\mathbb{R}^n} f(x) K_{\Omega,R}(x,y)\, dx. \end{aligned} \tag{47}$$

For the rest of the proof, We will put $K = K_{\Omega,R}$. Thus C has kernel K. Moreover,

$$\begin{aligned} \|K\|^2_{L^2(\mathbb{R}^n\times\mathbb{R}^n)} &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |\widetilde{\chi_\Omega}(y-x)|^2 \chi_R(x)^2\, dy\, dx \\ &= \int_{\mathbb{R}^n} \chi_R(x) \left(\int_{\mathbb{R}^n} |\widetilde{\chi_\Omega}(y-x)|^2\, dy \right) dx \\ &= |B_R|\ \|\widetilde{\chi_\Omega}\|_2^2 \overset{\text{Plancherel}}{=} |B_R|\ \|\chi_\Omega\|_2^2 = D_{\Omega,R} < \infty. \end{aligned} \tag{48}$$

Therefore C is Hilbert–Schmidt, and hence compact.

Seventh, the trace. Since C is compact and self-adjoint, the spectral theorem gives an orthonormal basis $(f_k)_{k\in\mathbb{N}}$ of PW_Ω formed by eigenvectors of C. We write $Cf_k = \lambda_k f_k$, with $\|f_k\|_2 = 1$. Since C is positive, $\lambda_k \geq 0$. Moreover, by (46),

$$\lambda_k = \langle Cf_k, f_k\rangle = \int_{B_R} |f_k(x)|^2\, dx.$$

In particular, $0 \leq \lambda_k \leq 1$. Since C is injective, no eigenvalue can be zero, and therefore $\lambda_k > 0$. Also $\lambda_k \neq 1$. Indeed, if $\lambda_k = 1$, then $\int_{B_R} |f_k(x)|^2\, dx = \|f_k\|_2^2$, so $f_k = 0$ almost everywhere on $\mathbb{R}^n \setminus B_R$. Since $f_k \in PW_\Omega$, it has a real analytic representative. Hence, as it vanishes on a non-empty open set, it must vanish identically, contradicting $\|f_k\|_2 = 1$. Thus $0 < \lambda_k < 1$.

We want to compute $\sum_k \lambda_k$. Before proceeding, We will fix a $x \in \mathbb{R}^n$ and define

$$\chi_{\Omega_x}(\xi) := e^{-2\pi i x\cdot\xi}\chi_\Omega(\xi).$$

Note that $\widetilde{\chi_{\Omega_x}}(y) = \widetilde{\chi_\Omega}(y-x)$. Also, we point out that $\widetilde{\chi_{\Omega_x}} \in PW_\Omega$, since $\widehat{\widetilde{\chi_{\Omega_x}}} = \chi_{\Omega_x}$, and clearly $\operatorname{supp}\chi_{\Omega_x} \subset \overline{B_\Omega}$. Moreover, $\widetilde{\chi_{\Omega_x}}$ represents evaluation at x on PW_Ω, since for every $f \in PW_\Omega$

$$\begin{aligned} \langle f, \widetilde{\chi_{\Omega_x}}\rangle &\overset{\text{Plancherel}}{=} \left\langle \widehat{f}, \chi_{\Omega_x} \right\rangle = \int_{\mathbb{R}^n} \widehat{f}(\xi)\overline{e^{-2\pi i x\cdot\xi}\chi_\Omega(\xi)}d\xi \\ &= \int_{B_\Omega} \widehat{f}(\xi) e^{2\pi i x\cdot\xi} d\xi = f(x). \end{aligned} \tag{49}$$

Now we are able to compute $\sum_k |f_k(x)|^2$ and $\sum_k \lambda_k$:

$$\sum_{k=1}^{\infty} |f_k(x)|^2 \overset{\text{Eq. (49)}}{=} \sum_{k=1}^{\infty} |\langle f_k, \widetilde{\chi_{\Omega_x}}\rangle|^2 = \sum_{k=1}^{\infty} |\langle \widetilde{\chi_{\Omega_x}}, f_k\rangle|^2 \overset{\text{Parseval}}{=} \|\widetilde{\chi_{\Omega_x}}\|_2^2 \overset{\text{Plancherel}}{=} \left\|\widehat{\widetilde{\chi_{\Omega_x}}}\right\|_2^2 = |B_\Omega|,$$

$$tr(C) := \sum_{k=1}^{\infty} \lambda_k = \sum_{k=1}^{\infty} \int_{B_R} |f_k(x)|^2\, dx \overset{\text{monotone convergence}}{=} \int_{B_R} \sum_{k=1}^{\infty} |f_k(x)|^2\, dx = \int_{B_R} |B_\Omega|\, dx = D_{\Omega,R}.$$

Eighth, the trace of C^2. When computing $\sum_k \lambda_k^2$, we see that the same argument used above gives a slightly more general identity. For fixed $x \in \mathbb{R}^n$, since $\widetilde{\chi_{\Omega_x}} \in PW_\Omega$, we have the following equality in $L^2(\mathbb{R}^n)$ in the variable y:

$$\begin{aligned} \widetilde{\chi_\Omega}(y-x) = \widetilde{\chi_{\Omega_x}}(y) &\overset{\text{ONB expansion}}{=} \sum_{k=1}^{\infty} \langle \widetilde{\chi_{\Omega_x}}, f_k\rangle\, f_k(y) \\ &= \sum_{k=1}^{\infty} \overline{\langle f_k, \widetilde{\chi_{\Omega_x}}\rangle} f_k(y) \overset{\text{Eq. (49)}}{=} \sum_{k=1}^{\infty} \overline{f_k(x)}\, f_k(y) = \sum_{k=1}^{\infty} f_k(y)\, \overline{f_k(x)}. \end{aligned} \tag{50}$$

Using $Cf_k = \lambda_k f_k$ and an intermediate step of (45) that reads $\langle Cf, g\rangle = \cdots = \langle f, \chi_R g\rangle$, we get, for every $N \in \mathbb{N}$

$$\begin{aligned} \sum_{k=1}^{N} \lambda_k^2 = \sum_{k=1}^{N} \langle C^2 f_k, f_k\rangle &\overset{\text{Eq.(45)}}{=} \sum_{k=1}^{N} \langle Cf_k, \chi_R f_k\rangle = \sum_{k=1}^{N} \int_{B_R} Cf_k(y)\overline{f_k(y)}\, dy \\ &\overset{\text{Eq. (47)}}{=} \sum_{k=1}^{N} \int_{B_R} \left(\int_{\mathbb{R}^n} \chi_R(x) \widetilde{\chi_\Omega}(y-x) f_k(x)\, dx \right) \overline{f_k(y)}\, dy \\ &= \int_{B_R} \int_{\mathbb{R}^n} \chi_R(x) \widetilde{\chi_\Omega}(y-x) \underbrace{\left(\sum_{k=1}^{N} f_k(x)\overline{f_k(y)} \right)}_{S_N^y(x)} dx\, dy. \end{aligned} \tag{51}$$

We want to take N to infinity, so let's check that we can commute both integrals with the limit in the summation. Since $\chi_R(\cdot)\widetilde{\chi_\Omega}(y-\cdot) \in L^2(\mathbb{R}^n)$, we may pass to the limit in the inner integral in (51):

$$\lim_{N\nearrow\infty} \int_{\mathbb{R}^n} \chi_R(x)\widetilde{\chi_\Omega}(y-x) S_N^y(x)\, dx \overset{L^2(\mathbb{R}^n)}{=} \int_{\mathbb{R}^n} \chi_R(x)\widetilde{\chi_\Omega}(y-x) \lim_{N\nearrow\infty} S_N^y(x)$$

Notice that by (50), for each fixed y

$$S_N^y(\cdot) = \sum_{k=1}^{N} f_k(\cdot)\overline{f_k(y)} \xrightarrow[N\nearrow\infty]{L^2(\mathbb{R}^n)} \widetilde{\chi_\Omega}(\cdot - y) = \widetilde{\chi_{\Omega_y}}(\cdot).$$

Moreover, by Cauchy–Schwarz applied to (51)

$$\left|\int_{\mathbb{R}^n} \chi_R(x)\widecheck{\chi_\Omega}(y-x)\left(\sum_{k=1}^N f_k(x)\overline{f_k(y)}\right)dx\right| \le \|\chi_R(\cdot)\widecheck{\chi_\Omega}(y-\cdot)\|_2 \left\|\sum_{k=1}^N f_k(\cdot)\overline{f_k(y)}\right\|_2$$
$$\le \|\widecheck{\chi_\Omega}\|_2 \left(\sum_{k=1}^\infty |f_k(y)|^2\right)^{1/2} = |B_\Omega|.$$

This bound is integrable on B_R allowing dominated convergence in the variable y. We can now take $N \nearrow \infty$ in (51)

$$\begin{aligned}\sum_{k=1}^\infty \lambda_k^2 = \lim_N \sum_{k=1}^N \lambda_k^2 &\overset{\text{dominated convergence}}{=} \int_{B_R}\int_{\mathbb{R}^n} \chi_R(x)\widecheck{\chi_\Omega}(y-x)\lim_N S_N^y(x)\,dx\,dy\\ &= \int_{B_R}\int_{\mathbb{R}^n} \chi_R(x)\widecheck{\chi_\Omega}(y-x)\widecheck{\chi_\Omega}(x-y)\,dx\,dy\\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |\widecheck{\chi_\Omega}(x-y)|^2\chi_R(x)\chi_R(y)dx\,dy.\end{aligned}$$

In the last equality we used that $\widecheck{\chi_\Omega}(-z) = \overline{\widecheck{\chi_\Omega}(z)}$, because χ_Ω is real-valued.

It remains to estimate this integral. We write $\sum_{k=1}^\infty \lambda_k^2 = I - J$, where

$$\begin{aligned} I &:= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |\widecheck{\chi_\Omega}(x-y)|^2\chi_R(x)\,dy\,dx,\\ J &:= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |\widecheck{\chi_\Omega}(x-y)|^2\chi_R(x)(1-\chi_R(y))\,dy\,dx.\end{aligned} \tag{52}$$

The first term was computed above in (48): $I = D_{\Omega,R}$. We now estimate J. We use the standard estimate of $\widecheck{\chi_\Omega}(z)$, which is a Bessel function, $|\widecheck{\chi_\Omega}(z)|^2 \lesssim \frac{\Omega^{2n}}{1+|\Omega z|^{n+1}}$ for $z \in \mathbb{R}^n$. We have the following inequality:

$$
\begin{aligned}
2J &\lesssim \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{\Omega^{2n}\chi_R(x)(1-\chi_R(y))}{1+|\Omega(x-y)|^{n+1}}\,dy\,dx = \int_{|x|<R}\int_{|y|>R} \frac{\Omega^{2n}\,dy\,dx}{1+|\Omega(x-y)|^{n+1}} \\
&\le \int_{|x|<R}\int_{|y-x|>R-|x|} \frac{\Omega^{2n}\,dy\,dx}{1+|\Omega(y-x)|^{n+1}} \quad \text{since } |y|>R \Rightarrow |y-x|\ge R-|x| \\
&\overset{u=y-x}{=\!=\!\downarrow} \int_{|x|<R} \Omega^{2n}\int_{|u|>R-|x|} \frac{du\,dx}{1+|\Omega u|^{n+1}} \overset{v=\Omega u}{=\!=\!\downarrow} \int_{|x|<R}\Omega^n\int_{|v|>\Omega(R-|x|)} \frac{dv\,dx}{1+|v|^{n+1}} \\
&\overset{*}{\lesssim} \Omega^n\int_{|x|<R}\frac{1}{1+\Omega(R-|x|)}\,dx \lesssim \Omega^n\int_0^R \frac{r^{n-1}}{1+\Omega(R-r)}\,dr \\
&\overset{t=\Omega r}{=\!=\!\downarrow} \Omega^n\int_0^{R\Omega}\frac{\left(\frac{t}{\Omega}\right)^{n-1}}{1+R\Omega-t}\frac{dt}{\Omega} = \int_0^{R\Omega}\frac{t^{n-1}}{1+R\Omega-t}\,dt \\
&\overset{0\le t\le R\Omega}{\le\!\downarrow} (R\Omega)^{n-1}\int_0^{R\Omega}\frac{1}{1+R\Omega-t}\,dt = (R\Omega)^{n-1}\left[-\log(1+R\Omega-t)\right]_0^{R\Omega} \\
&= (R\Omega)^{n-1}\log(1+R\Omega) = o\left((R\Omega)^n\right), \quad R\Omega\to\infty.
\end{aligned}
$$

In $*$ we have used the estimate $\int_{|v|>a}\frac{dv}{1+|v|^{n+1}} \lesssim \frac{1}{1+a}$ for $a\ge 0$, which follows from polar coordinates. If $0\le a\le 1$, then

$$
\int_{|v|>a}\frac{dv}{1+|v|^{n+1}} \le \int_{\mathbb{R}^n}\frac{dv}{1+|v|^{n+1}} \lesssim 1 \lesssim \frac{1}{1+a}.
$$

If $a>1$, then

$$
\int_{|v|>a}\frac{dv}{1+|v|^{n+1}} \lesssim \int_a^\infty \frac{r^{n-1}}{r^{n+1}}\,dr = \int_a^\infty r^{-2}\,dr = \frac{1}{a} \lesssim \frac{1}{1+a}.
$$

We conclude that

$$
tr(C^2) = \sum_{k=1}^\infty \lambda_k^2 = I - J = D_{\Omega,R} \pm o\left((R\Omega)^n\right), \qquad R\Omega\to\infty.
$$

□

Proposition A.2. *For any $0<t<1$:*

$$
D_{\Omega,R} - \frac{o(D_{\Omega,R})}{1-t} \le \#\{k : \lambda_k(\Omega,R) > t\} \le D_{\Omega,R} + \frac{o(D_{\Omega,R})}{t}, \qquad R\Omega\to\infty.
$$

This proposition says that, for every fixed threshold $0<t<1$, the number of eigenvalues above t is asymptotically equal to $D_{\Omega,R}$ when $R\Omega\to\infty$. In particular, if $0<\varepsilon<1$, then $\#\{k : \varepsilon < \lambda_k(\Omega,R) < 1-\varepsilon\} = o(D_{\Omega,R})$. Thus, apart from $o(D_{\Omega,R})$ intermediate eigenvalues, they lie near 0 or near 1. The ones close to 1 correspond to bandlimited functions which are well concentrated in B_R, whereas the ones close to 0 correspond to bandlimited functions whose mass is mostly outside B_R.

Proof. Set $\lambda_k := \lambda_k(\Omega, R)$ and $\mu = \mu_{\Omega,R} := \sum_{k=1}^{\infty} \delta_{\lambda_k}$.[40] Then $\mu((t,1)) = \#\{k : \lambda_k > t\}$. By A.1,

$$\int_0^1 x \, d\mu(x) = \sum_k \lambda_k = D_{\Omega,R}, \qquad \int_0^1 x^2 \, d\mu(x) = \sum_k \lambda_k^2 = D_{\Omega,R} - a_{\Omega,R},$$

where $a_{\Omega,R} := D_{\Omega,R} - \sum_k \lambda_k^2 = \sum_k (\lambda_k - \lambda_k^2) \geq 0$; $a_{\Omega,R} = o(D_{\Omega,R})$. Now define

$$q_t(x) := \frac{x(x-t)}{1-t}, \qquad p_t(x) := \frac{-x^2 + (1+t)x}{t}.$$

For $0 \leq x < 1$, we have $q_t(x) \leq \chi_{(t,1)}(x) \leq p_t(x)$. Indeed, if $0 \leq x \leq t$, then $q_t(x) \leq 0 \leq p_t(x)$. If $t < x < 1$, then

$$q_t(x) \leq 1 \iff (x-1)(x+1-t) \leq 0, \qquad 1 \leq p_t(x) \iff (x-t)(1-x) \geq 0.$$

Since $C_{\Omega,R}$ is injective, 1 is not an eigenvalue; hence the endpoint $x = 1$, where the first inequality would fail, gives no mass for μ. Therefore

$$\int_0^1 q_t \, d\mu \leq \mu((t,1)) \leq \int_0^1 p_t \, d\mu.$$

Using the two moment identities,

$$\int_0^1 q_t \, d\mu = \frac{1}{1-t} \int_0^1 x^2 \, d\mu - \frac{t}{1-t} \int_0^1 x \, d\mu = D_{\Omega,R} - \frac{a_{\Omega,R}}{1-t},$$
$$\int_0^1 p_t \, d\mu = -\frac{1}{t} \int_0^1 x^2 \, d\mu + \frac{1+t}{t} \int_0^1 x \, d\mu = D_{\Omega,R} + \frac{a_{\Omega,R}}{t}.$$

Hence

$$D_{\Omega,R} - \frac{a_{\Omega,R}}{1-t} \leq \#\{k : \lambda_k > t\} \leq D_{\Omega,R} + \frac{a_{\Omega,R}}{t}.$$

□

[40] The definition of the integral of an indicator function with respect to a measure says that, if $E \subset [0,1]$ is measurable and ν is a measure, then $\int_0^1 \chi_E(x) \, d\nu(x) = \nu(E)$. Applying this to the Dirac measure $\nu = \delta_a$, we obtain $\int_0^1 \chi_E(x) \, d\delta_a(x) = \delta_a(E)$. But, by definition of both,

$$\delta_a(E) = \begin{cases} 1, & a \in E, \\ 0, & a \notin E, \end{cases} \qquad \chi_E(a) = \begin{cases} 1, & a \in E, \\ 0, & a \notin E. \end{cases}$$

Therefore, $\int_0^1 \chi_E(x) \, d\delta_a(x) = \chi_E(a)$. By linearity, if φ is a simple function, say $\varphi = \sum_{j=1}^m c_j \chi_{E_j}$, then

$$\int_0^1 \varphi(x) \, d\delta_a(x) = \sum_{j=1}^m c_j \int_0^1 \chi_{E_j}(x) \, d\delta_a(x) = \sum_{j=1}^m c_j \chi_{E_j}(a) = \varphi(a).$$

Passing from simple to measurable functions, we get $\int_0^1 \varphi(x) \, d\delta_a(x) = \varphi(a)$. Finally, if $\mu = \sum_{k=1}^{\infty} \delta_{\lambda_k}$, then, we have

$$\int_0^1 \varphi(x) \, d\mu(x) = \int_0^1 \varphi(x) \, d\left(\sum_{k=1}^{\infty} \delta_{\lambda_k} \right)(x) = \sum_{k=1}^{\infty} \int_0^1 \varphi(x) \, d\delta_{\lambda_k}(x) = \sum_{k=1}^{\infty} \varphi(\lambda_k).$$

Bibliography

[ABR01] Sheldon Axler, Paul Bourdon, and Wade Ramey. *Harmonic Function Theory*, volume 137 of *Graduate Texts in Mathematics*. Springer, New York, 2 edition, 2001.

[Bea12] Jen Beatty. The radon transform and the mathematics of medical imaging. Honors thesis, Colby College, 2012. Honors Theses, Paper 646.

[Cer10] Joan Cerdà. *Linear Functional Analysis*, volume 116 of *Graduate Studies in Mathematics*. American Mathematical Society, 2010.

[CMOCP07] Carme Cascante, Xavier Massaneda, Joaquim Ortega-Cerdà, and Daniel Pascuas. Computerized tomography and the radon transform. *SeMA Journal: Boletín de la Sociedad Española de Matemática Aplicada*, (41):41–74, 2007.

[GJSM04] A. S. Granero, M. Jiménez-Sevilla, and J. P. Moreno. Intersections of closed balls and geometry of banach spaces. *Extracta Mathematicae*, 19(1):55–92, 2004.

[Grö01] Karlheinz Gröchenig. *Foundations of Time-Frequency Analysis*. Birkhäuser, 2001.

[GW18] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Pearson, 4 edition, 2018.

[HH15] John H. Hubbard and Barbara Burke Hubbard. *Vector Calculus, Linear Algebra, and Differential Forms: A Unified Approach*. Matrix Editions, Ithaca, NY, 5 edition, 2015.

[Nat01] Frank Natterer. *The Mathematics of Computerized Tomography*, volume 32 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, 2001. Unabridged republication of the 1986 edition.

[RS80] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics I: Functional Analysis*. Academic Press, revised and enlarged edition, 1980.

[Sil26] Luis Silvestre. Partial differential equations, May 2026. Lecture notes, University of Chicago.

[Str82] Robert S. Strichartz. Radon inversion—variations on a theme. *The American Mathematical Monthly*, 89(6):377–384, 1982.

[VG25] Diego Viedma Gordillo. Computerized tomography and the Radon transform through the lens of medical imaging. Treball final de grau, Universitat de Barcelona, Barcelona, 2025.