

# Math 5001 Homework 5

Jacob Hauck

October 10, 2022

## Question 1.

First, observe that

$$\frac{1}{2\pi} \int_0^{2\pi} \mathcal{R}f(x \cdot \omega, \omega) \, d\theta = \frac{1}{\pi} \int_0^\pi \mathcal{R}f(x \cdot \omega, \omega) \, d\theta \quad (1)$$

because, by the definition of  $\mathcal{R}$ ,

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} \mathcal{R}f(x \cdot \omega, \omega) \, d\theta &= \frac{1}{2\pi} \int_0^{2\pi} \int_{-\infty}^{\infty} f((x \cdot \omega)\omega + s\hat{\omega}) \, ds \, d\theta \\ &= \frac{1}{2\pi} \int_0^\pi \int_{-\infty}^{\infty} f((x \cdot \omega)\omega + s\hat{\omega}) \, ds \, d\theta + \frac{1}{2\pi} \int_\pi^{2\pi} \int_{-\infty}^{\infty} f((x \cdot \omega)\omega + s\hat{\omega}) \, ds \, d\theta \end{aligned}$$

Letting  $\varphi = \theta - \pi$  in the outer integral of the last double integral, and noting that  $\omega(\varphi) = -\omega(\theta)$ , and  $\hat{\omega}(\varphi) = -\hat{\omega}(\theta)$ , we get

$$\frac{1}{2\pi} \int_\pi^{2\pi} \int_{-\infty}^{\infty} f((x \cdot \omega)\omega + s\hat{\omega}) \, ds \, d\theta = \frac{1}{2\pi} \int_0^\pi \int_{-\infty}^{\infty} f((x \cdot \omega)\omega - s\hat{\omega}) \, ds \, d\theta$$

Finally, with the substitution  $s \rightarrow -s$  in the inner integral of the double integral, we get

$$\frac{1}{2\pi} \int_\pi^{2\pi} \int_{-\infty}^{\infty} f((x \cdot \omega)\omega + s\hat{\omega}) \, ds \, d\theta = \frac{1}{2\pi} \int_0^\pi \int_{-\infty}^{\infty} f((x \cdot \omega)\omega + s\hat{\omega}) \, ds \, d\theta$$

from which (1) follows. Now consider the change of variables  $y(s, \theta) = (x \cdot \omega)\omega + s\hat{\omega}$  in the integral

$$\frac{1}{\pi} \int_0^\pi \mathcal{R}f(x \cdot \omega, \omega) \, d\theta = \frac{1}{\pi} \int_0^\pi \int_{-\infty}^{\infty} f((x \cdot \omega)\omega + s\hat{\omega}) \, ds \, d\theta \quad (2)$$

This change of variables is valid because it is invertible a.e. if  $(s, \theta) \in \mathbb{R} \times [0, \pi)$ . For each  $\theta_0 \in [0, \pi)$ , the vector function  $y(\theta_0, s)$  describes a straight line with normal vector  $\omega(\theta_0)$ . Because  $0 \leq \theta_0 < \pi$ , each line points in a different direction. Furthermore, each line contains the point  $x$  because  $(x - (x \cdot \omega)\omega) \cdot \omega = 0$  (parallel projection of  $x$  onto the line is orthogonal to the normal iff  $x$  belongs to the line). Therefore, as long as  $y \neq x$ , there is a unique line containing  $y$  and  $x$ , which must correspond to a unique choice of  $(s, \theta)$  such that  $y = y(s, \theta)$ . This implies that  $y(s, \theta)$  is a.e. invertible.

The Jacobian of  $y(s, \theta)$  is given by

$$\begin{aligned} J &= \begin{pmatrix} \frac{\partial y_1}{\partial s} & \frac{\partial y_1}{\partial \theta} \\ \frac{\partial y_2}{\partial s} & \frac{\partial y_2}{\partial \theta} \end{pmatrix} = \begin{pmatrix} -\sin(\theta) & \frac{d}{d\theta}((x \cdot \omega) \cos(\theta)) - s \cos(\theta) \\ \cos(\theta) & \frac{d}{d\theta}((x \cdot \omega) \sin(\theta)) - s \sin(\theta) \end{pmatrix} \\ &= \begin{pmatrix} -\sin(\theta) & x \cdot \hat{\omega} \cos(\theta) - (x \cdot \omega) \sin(\theta) - s \cos(\theta) \\ \cos(\theta) & x \cdot \hat{\omega} \sin(\theta) + (x \cdot \omega) \cos(\theta) - s \sin(\theta) \end{pmatrix} \end{aligned}$$

because  $\frac{d\omega}{d\theta} = \hat{\omega}$ . Therefore  $|\det(J)| = |s - x \cdot \hat{\omega}|$ . Clearly  $s = y \cdot \hat{\omega}$ , so  $|\det(J)| = |(y - x) \cdot \hat{\omega}| = |y - x|$ . Applying the change of variables to (2) gives

$$\frac{1}{\pi} \int_0^\pi \mathcal{R}f(x \cdot \omega, \omega) d\theta = \frac{1}{\pi} \int_{\mathbb{R}^2} f(y) |\det(J)|^{-1} dy = \frac{1}{\pi} \int_{\mathbb{R}^2} \frac{f(y)}{|y - x|} dy$$

This is just  $\left(f * \frac{1}{|\cdot|}\right)(x)$ , which completes the demonstration.

If one used the integral in (1) as a reconstruction of  $f$ , one would not get  $f$ , but a blurred version of  $f$  because  $\int_{\mathbb{R}^2} \frac{f(y)}{|y - x|} dy$  is essentially an average of the values of  $f$  near  $x$ . For  $y$  far away from  $x$ , the convolution kernel  $\frac{1}{|y - x|}$  is small, and for  $y$  close to  $x$ , it is large. Thus, instead of  $f(x)$ , one gets a value close to  $f(x)$  averaged out with nearby values of  $f$ .

## Question 2.

A sound in a space  $R$  is defined by the (relative) pressure field  $p(t, x)$  throughout that space and over time. To the human ear, however, a sound is what happens when the varying pressure near the ear drum causes the membrane to vibrate; hearing is the act of recording and interpreting these vibrations. Thus, human hearing is a matter of recording the pressure field over time at a fixed location in space (where the ear drum is located).

Let a sound  $S$  occur with pressure field  $p(t, x)$ . A playback device causes a disturbance that creates a pressure field  $\tilde{p}(t, x)$ . The purpose of audio playback is to allow a *human ear* located at  $x = x_2$  to hear what *would have been* heard if the ear were listening to  $S$  at  $x = x_1$ ; thus, the device must ensure that  $\tilde{p}(t - t_0, x_2) = p(t, x_1)$ , where we allow a time shift of  $t_0$  because the playback sound is occurring at a later time.

Note that  $\tilde{p}$  may be as different from  $p$  as we like, as long as their temporal profiles at  $x_1$  and  $x_2$  match. In other words, a playback device need not construct the original sound  $S$ ; rather, it only needs to reconstruct what a human would have heard. For this reason, a recording device need only record a single temporal profile at  $x = x_1$ , and a playback device should require no more information to achieve its purpose.

Fortunately, it is practically simple to arrange for a playback device with the above property, as well as some even better properties. The strategy is to vibrate a small object in such a way that (roughly) spherical sound waves are produced with the property that at any point on some fixed sphere around the vibrating object the temporal pressure profile matches the recorded pressure profile.

If the playback device is used in empty space (empty except for air, of course), then the resulting pressure field will not only be spherically symmetrical, but temporally self-similar, in the sense that the temporal profiles at points two concentric spheres centered at the vibrating object will be the same up to a constant of proportionality and a temporal shift.

Having discussed properties of playback and recording devices, we are ready to describe the situation in “I am sitting in a room” mathematically.

- Let a recording device be located at  $x = x_{\mathcal{R}}$  in the room  $R$ . For a sound with pressure field  $p(t, x)$ , let the **recording operator**  $\mathcal{R}$  be defined by

$$\mathcal{R}p(t) = p(t, x_{\mathcal{R}})$$

so that  $r(t) = \mathcal{R}p(t)$  is the recording data.

- Assume we have a playback device capable of producing a spherically symmetrical pressure field  $\tilde{p}(t, x)$  from recording data  $r(t)$  such that  $\tilde{p}(t - t_0, x_{\mathcal{P}}) = r(t)$  at some point  $x = x_{\mathcal{P}}$  and for some time offset  $t_0$ . Let the **playback operator**  $\mathcal{P}$  be defined by

$$\mathcal{P}r(t, x) = \tilde{p}(t, x)$$

Then both  $\mathcal{R}$  and  $\mathcal{P}$  are shift-invariant (in time) and linear. For  $\mathcal{R}$ , this is an obvious consequence of the definition. For  $\mathcal{P}$ , it takes some reasoning.

Since  $\mathcal{P}r(t - t_0, x_{\mathcal{R}}) = r(t)$ , we know that  $\mathcal{P}r(t, x) = r(t + t_0)$  for all  $x$  in some sphere containing  $x_{\mathcal{R}}$  by the assumption of spherical symmetry. Due to the fact that sound waves obey the wave equation, the rest of  $\tilde{p}$  is uniquely determined by solving the wave equation with the spherical boundary condition derived above and zero initial pressure displacement and velocity. Solving the wave equation under this condition is shift invariant in time and linear. Therefore  $\mathcal{P}$  is shift-invariant in time and linear.

Let  $p_0(t, x)$  be the pressure field of a sound in empty space. Let  $R$  be a space containing objects, and suppose there exists  $t_0$  such that the support of  $p_0(t, x)|_{t < t_0}$  does not intersect any objects in the space. Then we can define the distorted pressure field  $p(t, x)$  to be the field obtained by solving the wave equation with reflection/absorption boundary conditions for the objects in the space and the initial condition that  $p(t, x) = p_0(t, x)$  for  $t < t_0$ . Let  $\mathcal{D}$  be the distortion operator for the space  $R$ , which is defined by

$$\mathcal{D}p_0(t, x) = p(t, x)$$

In the “I am sitting in a room” situation, both the initial sound and the playback sounds clearly satisfy the support requirement above, so we can use  $\mathcal{D}$  to find their distortions, which are what the recording device actually records. Thus, if  $p_0(t, x)$  is the pressure field that would have occurred due to the voice sound in an empty space, then the first recording is  $r_1(t) = \mathcal{R}\mathcal{D}p_0(t)$ . Following this, the pressure field that would be produced by the first playback in empty space is  $p_1(t, x) = \mathcal{P}r_1(t, x)$ . Therefore, the second recording is  $r_2(t) = \mathcal{R}\mathcal{D}p_1(t)$ .

More generally, if we assume that the voice is essentially a playback of a recording  $r_0(t)$  (this is not necessary, but it is convenient and more or less true), then we can write the  $n$ th recording  $r_n(t)$  as

$$r_n(t) = (\mathcal{R}\mathcal{D}\mathcal{P})^n r_0(t) = T^n r_0(t) \quad (3)$$

where  $T := \mathcal{R}\mathcal{D}\mathcal{P}$ .

If  $\mathcal{D}$  is shift-invariant (in time) and linear, then so, too, is  $T$  (because  $\mathcal{R}$  and  $\mathcal{P}$  are, as well). It is clear enough that  $\mathcal{D}$  should be shift-invariant in time; this follows from the shift-invariance of solving the wave equation (and simple intuition: if a sound is made at a later time in an unchanged room, then the distortion should be the same but occur at a correspondingly later time).

Linearity is trickier. I won't give a full proof, just a motivational case showing the linearity of  $\mathcal{D}$  in the case of an initial pressure field  $p_0$  resulting from a playback and a single planar surface reflecting the sound with attenuation. It seems plausible to build up linearity of an arbitrary room from this fact...

To this end, let  $S$  be a planar surface disjoint from the origin, and let  $p_0(t, x)$  be a pressure field of a sound in an empty space resulting from a playback centered at the origin. Suppose that the surface  $S$  reflects sound with attenuation coefficient  $a \in (0, 1)$ . Then applying the law of reflection to find  $\mathcal{D}p_0(t, x)$ , we see that

$$\mathcal{D}p_0(t, x) = p_0(t, x) + ap_0(t, x - x_r)\chi_K(x)$$

where  $x_r$  is the reflection of the origin over the plane containing  $S$ , and  $K$  is the set formed by the union of all rays that start at some point  $x \in S$  and emanate in the direction  $x - x_r$ . Clearly, in this case,  $\mathcal{D}$  is shift-invariant in time and linear.

Assuming that  $T$  is in fact shift-invariant and linear, we see that it is actually a Fourier multiplier operator. Let  $m(\tau)$  be its symbol. Assume that  $r_0 \in L_2$ . Then taking the Fourier transform of  $r_n(t)$  with respect to time on both sides of (3) we get

$$\hat{r}_n(\tau) = m^n(\tau)\hat{r}_0(\tau)$$

Let  $Z = \{\tau \mid |m(\tau)| < 1\}$ , and let  $I = \{\tau \mid |m(\tau)| > 1\}$ . If  $\tau \in Z$ , then  $m^n(\tau) \rightarrow 0$ ; if  $\tau \in I$ , then  $|m^n(\tau)| \rightarrow \infty$ . Otherwise,  $m^n(\tau)$  likely does not converge, but  $|m^n(\tau)| \rightarrow 0$ .

If  $\tau \in Z$ , then obviously  $\widehat{r}_n(\tau) \rightarrow 0$ , so some frequencies in the recordings are vanishing. If  $\tau \in I$ , then clearly  $|\widehat{r}_n(\tau)| \rightarrow \infty$ , so these frequencies start to account for all of the energy. In fact, the energy in the frequencies in  $Z$  after  $n$  iterations is

$$E_n \propto \int_Z |\widehat{r}_n(\tau)|^2 d\tau \rightarrow 0$$

as  $n \rightarrow \infty$  by dominated convergence. Then the energy in the other frequencies approaches the total energy. If the signal  $r(t)$  is played back after being scaled to a fixed volume, then, what one hears is just the frequencies in  $I$ . If  $I$  is a small (discrete) set, then one hears a sum of sine-waves at the frequencies in  $I$ . A sine-wave shaped pressure field sounds like a tuning-fork, and a sum such sounds gives something like the ethereal ringing experienced at the end of “I am sitting in a room” ( $n = 32$ ).

Attached is a sound file I made with 8 sine-waves at different frequencies playing at once (hopefully you can



actually open it).