1 Theory

1.1 Self-consistency of the Fourier series

Observation. The following is a valid representation of the Dirac delta function:

$$\sum_{k=-\infty}^{\infty} e^{iky} = 2\pi\delta(y).$$

Lemma. $\delta(ay) = \delta(y)/a$.

Proof. We start with a change of variables $y \mapsto y/a$, which leaves the limits of integraiton unchanged:

$$\int_{-\infty}^{\infty} f(y)\delta(ay)\,dy = \int_{-\infty}^{\infty} f(y/a)\delta(y)\,d(y/a) = \frac{1}{a}\int_{-\infty}^{\infty} f(y/a)\delta(y)\,dy = \frac{f(0)}{a} = \int_{-\infty}^{\infty} f(y)\left[\frac{\delta(y)}{a}\right] dy.$$

□

Claim. The Fourier series, as defined in equations (3) and (2) in the notes, is self-consistent.

Proof. Using the definition of $\tilde{h}_k$ and $f_k \equiv k/L$,

$$\begin{aligned}
h(x) &= \sum_{k=-\infty}^{\infty}\left[\frac{1}{L}\int_0^L h(x')e^{2\pi ikx'/L}\,dx'\right]e^{-2\pi ikx/L} \\
&= \int_0^L h(x')\left[\frac{1}{L}\sum_{k=-\infty}^{\infty} e^{2\pi ik(x'-x)/L}\right]dx' \\
&= \int_0^L h(x')\delta(x'-x)\,dx' = h(x).
\end{aligned}$$

□

1.2 Linear combination of exponentials

Lemma. $\sin(\theta+\phi) = \sin\theta\cos\phi + \cos\theta\sin\phi$.

Proof. Using Euler's formula:

$$\begin{aligned}
e^{i(\theta+\phi)} &= e^{i\theta}e^{i\phi} \\
\cos(\theta+\phi) + i\sin(\theta+\phi) &= (\cos\theta + i\sin\theta)(\cos\phi + i\sin\phi) \\
&= \cos\theta\cos\phi - \sin\theta\sin\phi + i(\sin\theta\cos\phi + \cos\theta\sin\phi).
\end{aligned}$$

Equating imaginary parts yields the desired result. □

Claim. $A\sin(2\pi x/L + \phi)$ is a linear combination of $e^{-2\pi ix/L}$ and $e^{2\pi ix/L}$ over the scalar field $\mathbb{C}$.

Proof. We use the trigonometric identity above and the definitions of sin and cos in terms of complex exponentials (from Euler's formula):

$$\begin{aligned}
A\sin(2\pi x/L + \phi) &= (A\cos\phi)\sin(2\pi x/L) + (A\sin\phi)\cos(2\pi x/L) \\
&= (A\cos\phi)\left(\frac{e^{2\pi ix/L} - e^{-2\pi ix/L}}{2i}\right) + (A\sin\phi)\left(\frac{e^{2\pi ix/L} + e^{-2\pi ix/L}}{2}\right) \\
&= \left[\frac{A(\sin\phi - i\cos\phi)}{2}\right]e^{2\pi ix/L} + \left[\frac{A(\sin\phi + i\cos\phi)}{2}\right]e^{-2\pi ix/L}.
\end{aligned}$$

□

1.3 Redundancy in Fourier coefficients of real functions

Claim. For $h(x) \in \mathbb{R}$, the Fourier coefficients $\tilde{h}_k$ satisfy $\tilde{h}_{-k} = \tilde{h}_k^*$.

Proof. Conjugation is linear: for $A, B \in \mathbb{R}$ and $\alpha, \beta \in \mathbb{C}$, $(A\alpha + B\beta)^* = A\alpha^* + B\beta^*$. Therefore, to show that two integral expressions are conjugates, it suffices to show that their integrands are conjugates:

$$\begin{aligned}
\tilde{h}_{-k} &= \frac{1}{L}\int_0^L h(x)e^{-2\pi i k x/L}\,dx \\
&= \frac{1}{L}\int_0^L h(x)\left(e^{2\pi i k x/L}\right)^*\,dx \\
&= \frac{1}{L}\int_0^L \left[h(x)e^{2\pi i k x/L}\right]^*\,dx \\
&= \left[\frac{1}{L}\int_0^L h(x)e^{2\pi i k x/L}\,dx\right]^* = \tilde{h}_k^*.
\end{aligned}$$

□

1.4 Convolution theorem

Claim. The Fourier coefficients of the product $H(x) = h^{(1)}(x)h^{(2)}(x)$ are given by the convolution:

$$\tilde{H}_k = \sum_{k'=-\infty}^{\infty} \tilde{h}^{(1)}_{k-k'}\tilde{h}^{(2)}_{k'}.$$

Proof. We express $h^{(1)}(x)$ and $h^{(2)}(x)$ as Fourier series and find the Cauchy product $H(x)$:

$$\begin{aligned}
H(x) &= \left[\sum_{k=-\infty}^{\infty} \tilde{h}^{(1)}_k e^{-2\pi i k x/L}\right]\left[\sum_{k'=-\infty}^{\infty} \tilde{h}^{(2)}_{k'} e^{-2\pi i k' x/L}\right] \\
&= \sum_{k=-\infty}^{\infty}\sum_{k'=-\infty}^{\infty} \left[\tilde{h}^{(2)}_{k'} e^{-2\pi i k' x/L}\right]\left[\tilde{h}^{(1)}_{k-k'} e^{-2\pi i (k-k') x/L}\right] \\
&= \sum_{k=-\infty}^{\infty}\sum_{k'=-\infty}^{\infty} \tilde{h}^{(1)}_{k-k'}\tilde{h}^{(2)}_{k'} e^{-2\pi i k x/L} \\
&= \sum_{k=-\infty}^{\infty} \tilde{H}_k e^{-2\pi i k x/L},
\end{aligned}$$

Equating the cofficients of the sum over k, we have the desired result. □

Graphically, the convolution of $\tilde{h}^{(1)}$ and $\tilde{h}^{(2)}$ can be interpreted as follows: pick $\tilde{h}^{(1)}_{k'}$ without loss of generality, since convolution is commutative. Flip the other function on the k' axis and add an offset: $\tilde{h}^{(2)}_{k-k'}$. As we integrate/sum over k', these two functions are moved together on the same axis – as if one cross through the other – and the convolution is the integral/sum of their product at each point.

1.5 Testing the `numpy` FFT

We compare analytical and numerical methods of obtaining the Fourier series of the following two functions:

$$g(t) = A\cos(ft + \varphi) + C \quad \text{and} \quad h(t) = A\exp\left[-B(t - L/2)^2\right].$$

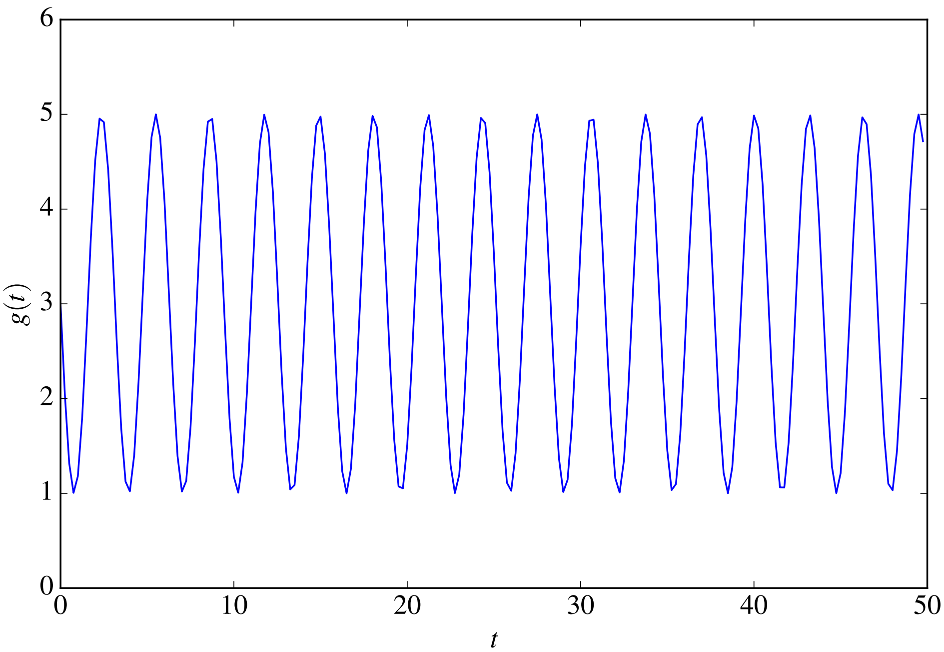

Figure 1: $g(t)$ vs. t with paramaters $A = 2$, $f = 2$, $\varphi = \pi/2$, and $C = 0$.

1.5.1 Cosine

Since $\cos t$ has period 2π, $g(t)$ has period $2\pi/f \equiv L$. The Fourier coefficients are:

$$\begin{aligned}
\tilde{g}_k &= \frac{1}{L}\int_0^L [A\cos(ft+\varphi)+C]e^{2\pi ikt/L}\,dt \\
&= \frac{f}{2\pi}\int_0^{2\pi/f} [A\cos(ft+\varphi)+C]e^{ikft}\,dt \\
&= \frac{f}{2\pi}\left[A\int_0^{2\pi/f}\cos(ft+\varphi)e^{ikft}\,dt + C\int_0^{2\pi/f} e^{ikft}\,dt\right] = \begin{cases} C & k=0 \\ \frac{A}{2}e^{-i\varphi} & k=\pm 1 \\ 0 & k\neq 0,\pm 1. \end{cases}
\end{aligned}$$

The first integrand is a product of two waves – a cosine with frequency f and a complex exponential with frequency kf – over full periods of both waves. They will destructively interfere everywhere except when $k = \pm 1$, creating matching frequencies. The second integral evaluates to $f/2\pi$ at $k = 0$, since the integrand is unity, and to 0 everywhere else. We expect isolated peaks at $k = \pm 1$ and a larger peak at $k = 0$. The FFT implementation in `numpy` makes it easy to convert the k-axis to a frequency axis (cycles per second, not angular), whereas f as given is an angular frequency, so we should see the isolated peaks at $\bar{f} = 0, \pm f/2\pi$.

A key difference between our analytical solution and the numerical implementation is that here, we chose L to be the period of the cosine. In `numpy`, L will be the length of our dataset. As long as L is large enough to capture a full oscillation, we should see the same peaks in the transform, but their height will scale with L as the integration interval will be larger.

Figure 1 shows $g(t)$ and figure 2 shows its transform. As expected, there are two isolated peaks in the Fourier transform $\tilde{g}(\bar{f})$ at $\bar{f} = \pm f/2\pi \approx \pm 0.32$ and a tall peak at $\bar{f} = 0$. The width of the peaks at $\bar{f} = \pm f/2\pi$ is due to the sampling – the input data from $g(t)$ does not perfectly trace a cosine. Inverse-transforming the data in figure 2 returns figure 1 as expected.

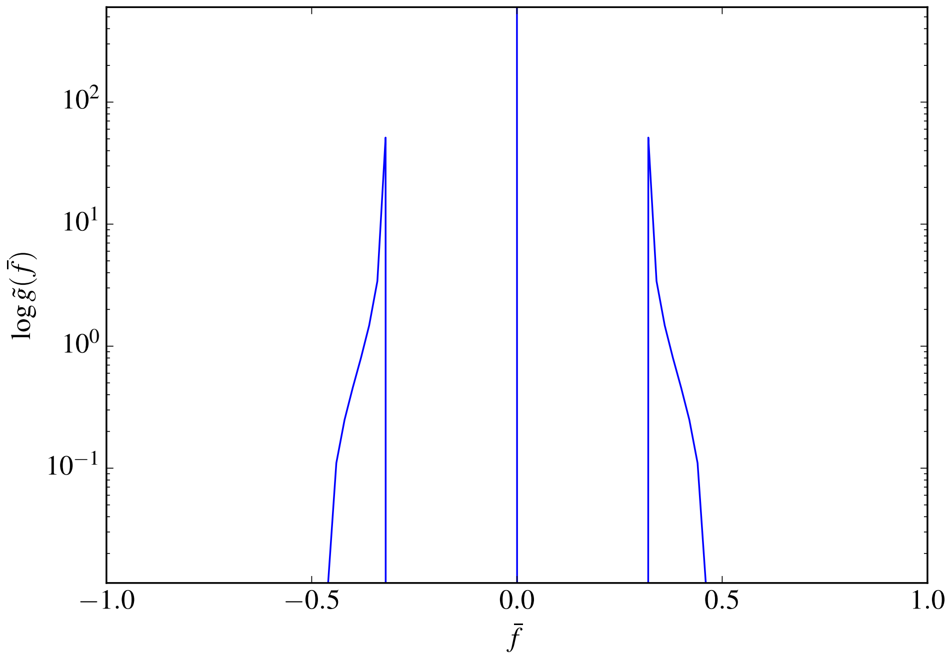

Figure 2: The Fourier transform $\tilde{g}(\bar{f})$ of the cosine vs. $\bar{f}$.

1.5.2 Gaussian

We are given that $h(t)$ is L-periodic for small enough B – that is, for small enough standard deviation, we can repeat Gaussians centered at multiples of the mean without the distributions bleeding into each other significantly. The Fourier coefficients of the Gaussian are given by:

$$\begin{aligned}\tilde{h}_k &= \frac{1}{L}\int_0^L Ae^{-B(t-L/2)^2}e^{2\pi ikt/L}\,dt \\ &= \frac{A\sqrt{\pi}}{2L\sqrt{B}}\left(\operatorname{erf}\gamma_+ + \operatorname{erf}\gamma_-\right)e^{i\pi k}\exp\left(-\frac{\pi^2k^2}{BL^2}\right) \quad \text{where} \quad \gamma_\pm \equiv \frac{BL^2 \pm 2\pi ik}{2L\sqrt{B}},\end{aligned}$$

so we expect oscillations due to the $e^{i\pi k}$ factor in a Gaussian envelope centered at $k = 0$. Figure 3 shows the Gaussian, and figure 4 shows its Fourier transform. The shapes are as expected, and applying the transform and inverse transform in succession returns figure 3.

2 Uniformly-sampled Arecibo data

2.1 Isolating the signal frequency

The Arecibo dataset is sampled at 1 ms intervals, so `numpy.fft.fftfreq` will return an f-axis in cycles per millisecond, or thousands of cycles per second (kHz). We multiply by 10^{-3} to get our signal in MHz.

Figure 5 shows the transform of the Arecibo data. A single peak, isolated from the noise, is visible. Figure 6 shows a closer view of this peak. At this scale, the finite width of the peak is visible. The peak is maximized at $|f - 1420\text{ MHz}| = 1.37 \times 10^{-4}$ MHz, so we detect a signal at $1420 \pm 1.37 \times 10^{-4}$ MHz. Here, the $\pm$ does not denote an uncertainty, but is a result of the symmetry of the real-valued Fourier transform.

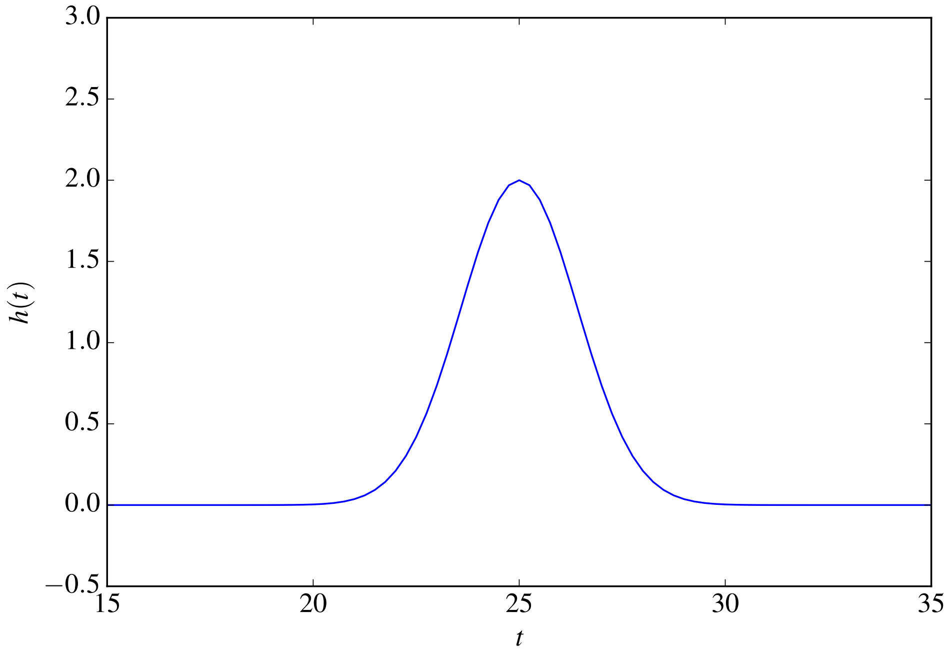

Figure 3: $h(t)$ vs. t with parameters $L = 50$ and $B = 0.25$.

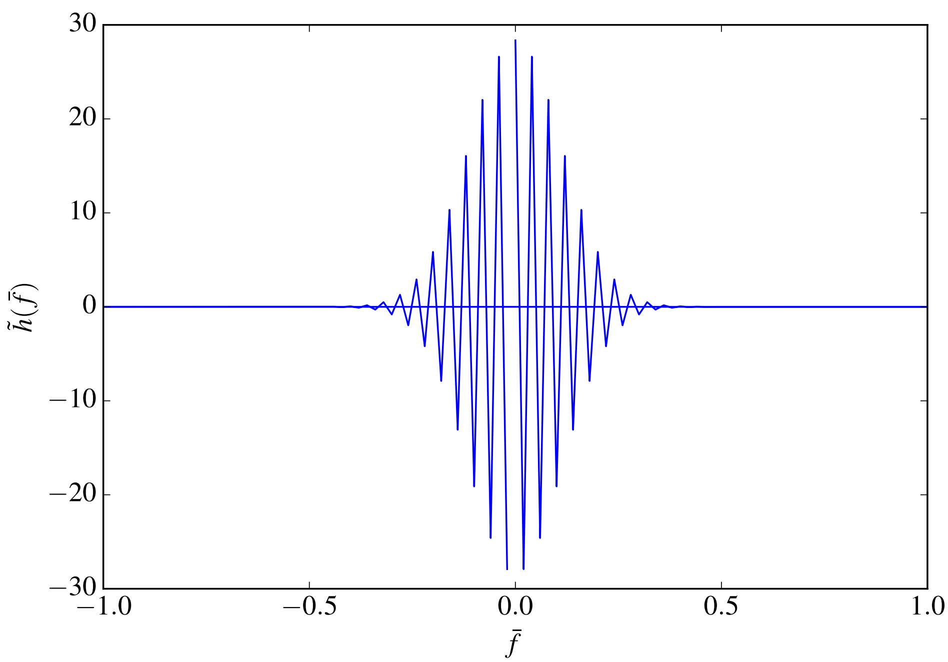

Figure 4: The Fourier transform $\tilde{h}(\bar{f})$ of the Gaussian vs. $\bar{f}$.

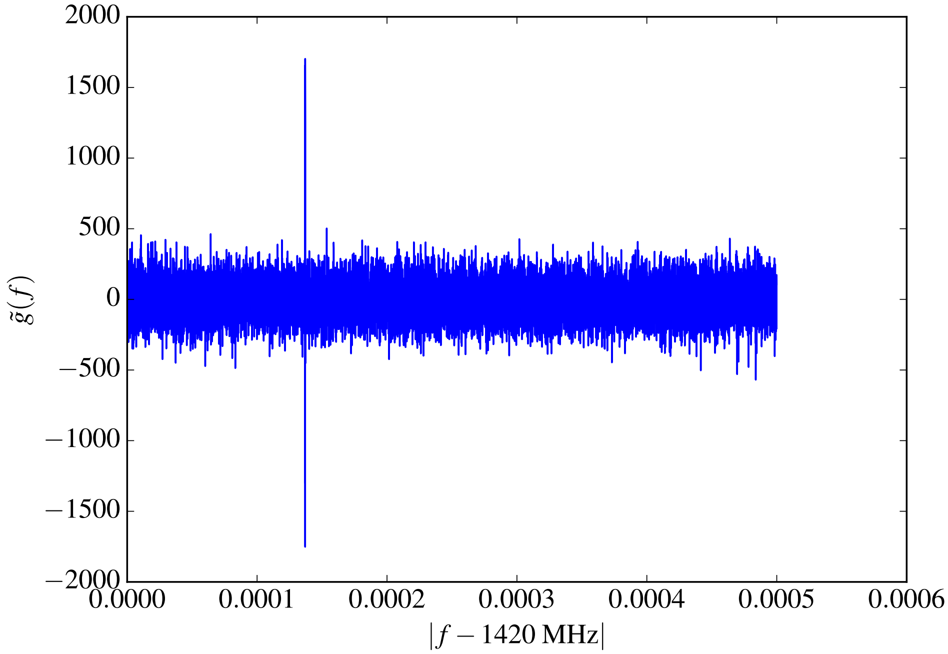

Figure 5: Fourier transform of the Arecibo data. Frequencies [MHz] on the x-scale have been shifted such that zero corresponds to 1420 MHz.

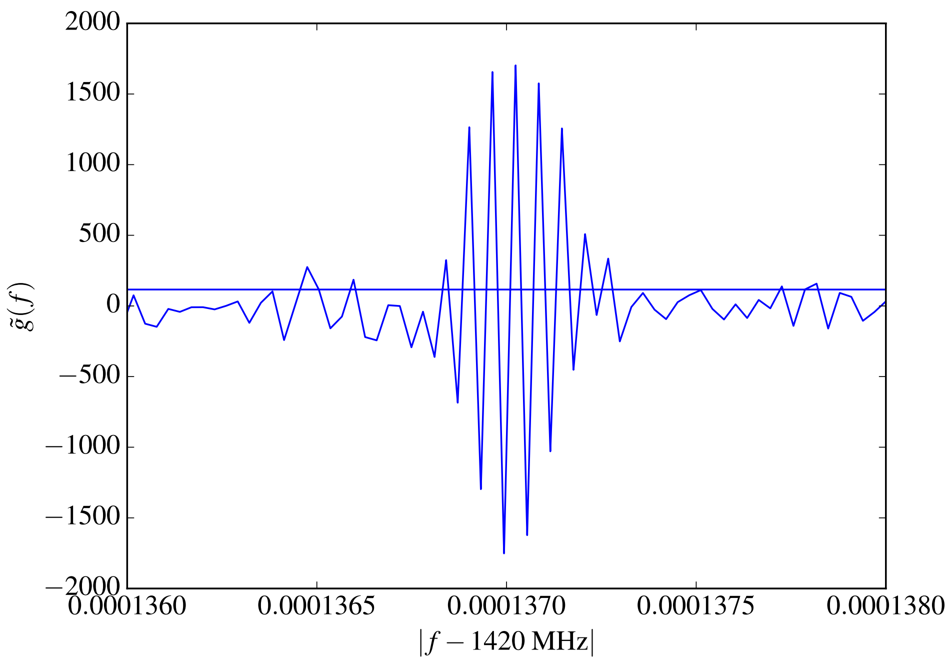

Figure 6: Closer view of the single peak visible in figure 5.

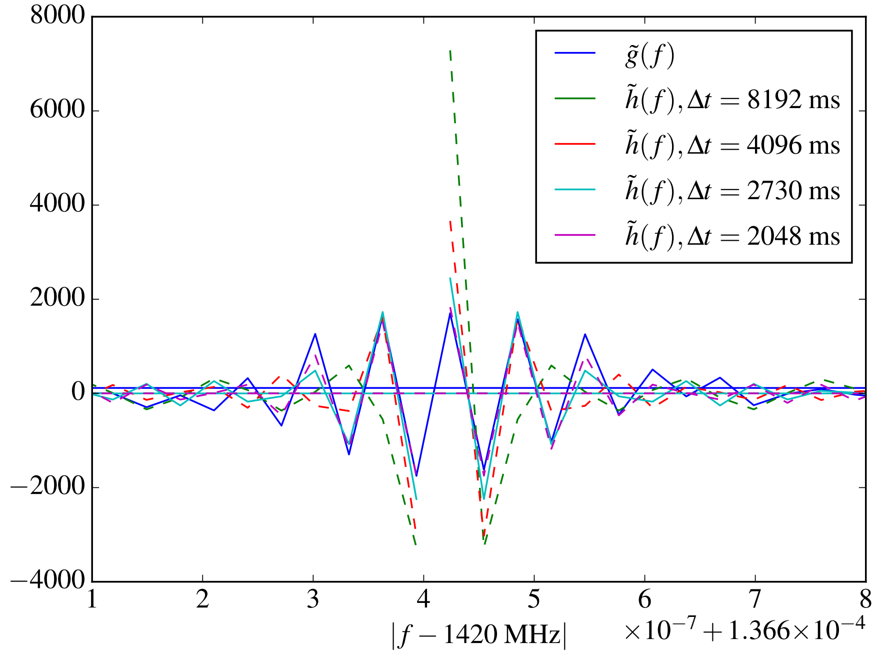

Figure 7: Comparison of the signal transform, $\tilde{g}(f)$, with the transforms of Gaussian envelopes, $\tilde{h}(f)$, with various widths Δt.

2.2 Gaussian envelope

If the signal is a perfect sinusoid multiplied by a Gaussian envelope, then we expect the Fourier transform of the signal to be the convolution of a delta function (transform of the sinusoid) and a Gaussian-enveloped wave (transform of the Gaussian as in figure 4).

Consider the graphical interpretation of convolution as explained in section 1.4. If we hold the Gaussian transform and move a delta function along the same axis, we note that the product of the two functions will be nonzero only as the delta function (which is nonzero at an isolated location) is moving through the Gaussian transform. The resulting convolution will therefore have the same width as the Gaussian transform.

Our Gaussian is of the form:

$$h(t) = \exp\left[-\frac{(t-t_0)^2}{\Delta t^2}\right].$$

To determine Δt, we plot several transform $\tilde{h}(f)$ generated from different Δt and compare the transform width to the signal. We saw in section 1.5.2 that a Gaussian with mean $L/2$ yields a transform centered around $k = 0$, so we will set $t_0 = (32768/2)$ ms and shift the resulting transform to $|f - 1420\ \text{MHz}| = 1.37 \times 10^{-4}$ MHz, the central frequency of the signal.

Figure 7 shows this comparison for $\Delta t = \frac{t_0}{2}, \frac{t_0}{4}, \frac{t_0}{6}, \frac{t_0}{8}$. Of these, the Gaussian transform for $\Delta t = \frac{t_0}{4} = 4096$ ms matches the signal transform width best.

3 Lomb-Scargle routine for unequally sampled data

3.1 Lomb-Scargle implementation from `scipy`

We use the implementation of the Lomb-Scargle algorithm provided by `scipy.signal.lombscargle`.

3.2 Testing Lomb-Scargle on evenly-sampled data

Using Lomb-Scargle on evenly-sampled data reveals a documented bug in the `scipy` implementation. According to a bug report[1], the `scipy` implementation calculates the arctangent of a ratio whose denominator is zero for evenly-sampled data.

This issue was fixed in `scipy` version 0.15.0, but my distribution provides version 0.14.1. Since the only reason to use Lomb-Scargle over the FFT in practice is to study unevenly-sampled data, I tried to simulate unevenly-sampled data from the evenly sampled Arecibo data by randomly dropping 10% of the data and by randomly shifting each time by a small amount. However, the `ZeroDivisionError` persists, even with real unevenly-sampled data (from CRTS).

[1]available at `https://github.com/scipy/scipy/issues/3787`