Tre DiPassio's Solution to the 2019 International Student Challenge Problem in Acoustic Signal Processing

Joseph (Tre) DiPassio III
Electrical and Computer Engineering, University of Rochester, Rochester, NY 14627, USA

In this year's Student Challenge Problem, a turboprop aircraft is flying above a body of water and is being recorded by a hydrophone at some depth beneath the surface[1]. The plane's speed, height, and propeller frequency all influence the sound field at the hydrophone's position. In his 1972 work, R.J. Urick proposed four different acoustic paths from a source in the air medium to a submerged receiver that influence the underwater sound field[2]. In particular, we are asked to consider the direct refraction path, and to assume the speed of sound in both media is constant. For this direct path between two iso-speed media, Urick detailed a geometric approach to placing a virtual source in the water medium to model the true source in the air medium.

This solution explores a computational approach to placing this virtual source in the water medium, and using its trajectory to create a predictive model for the instantaneous frequency observed at the hydrophone position by using Doppler effect calculations. The accuracy of this model will be evaluated via comparison with given experimental data. The model will then be used to extract the speed, height, and propeller frequency of the aircraft source for a given hydrophone recording. `(Animated versions of select figures and a verbal presentation of this solution are posted on TreDiPassio.com/ASA2019)`

I. BACKGROUND AND THEORY

A. Snell's Law

Snell's Law in acoustics describes how an acoustic ray refracts through a boundary of two media[3]. In particular, Snell's Law states that the ratio of the sines of the refracted and incident angles relative to the normal are proportional to the ratio of the speeds of sound in the media. This relationship is given by

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{c_1}{c_2} \tag{1}$$

where c_1 and c_2 are the speeds of sound and θ_1 and θ_2 are the angles with respect to the normal of the rays in their respective medium. The critical angle of incidence, θ_c, is defined as the angle incident to the boundary such that the refracted angle $\theta_2 = 90°$, thus $\sin\theta_2 = 1$. At this refracted angle, all acoustic energy will radiate along the surface of the boundary. For an air-water boundary, $c_1 = 340\ {}^{m}/_{s}$ and $c_2 = 1520\ {}^{m}/_{s}$, therefore

$$\theta_c = \arcsin(\frac{c_1}{c_2}) \approx 13° \tag{2}$$

Thus, for the given air-water boundary, only incident angles between 0° and roughly 13° will be considered.

There will be one angle of incidence in this range from the acoustic source (in this case the aircraft) to the air-ocean boundary that both satisfies Snell's Law and that achieves the correct angle of refraction to exactly reach the hydrophone's relative horizontal position and depth. The incident angle corresponding to this so-called direct sound path can be found using an iterative algorithm described below. Firstly, the situation's geometry can be described via the conventions in Figure 1.

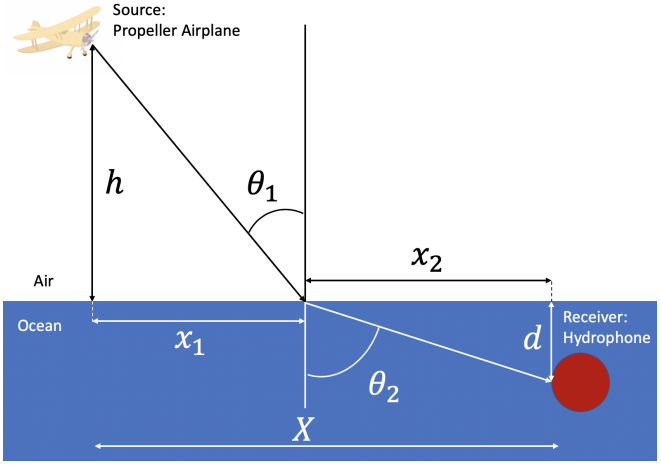


FIG. 1. Plane-Hydrophone direct path geometry. h is the plane's height above the ocean surface, and d is the hydrophone's depth beneath the surface. x_1 is the horizontal distance from the plane to the direct sound path's crossing point through the boundary, x_2 is the horizontal distance from this crossing point to the hydrophone's position, and X is the total horizontal distance between source and receiver.

From this geometry, the following system of equations can be created:

$$x_1 = h \tan\theta_1 \tag{3a}$$
$$x_2 = d \tan\theta_2 \tag{3b}$$
$$X = x_1 + x_2, \tag{3c}$$

For any point in the plane's trajectory, horizontal distance X is known, and θ_2 can be written in terms of θ_1 using Snell's Law from Equation 1. Therefore, θ_1 and X are related with the following substitution:

$$X = h \tan\theta_1 + d \tan(\arcsin(\frac{c_2}{c_1} \sin\theta_1)) \tag{4}$$

As previously stated, there is one and only one value of θ_1 in the appropriate range that will satisfy Equation 4 for the value of X determined by the plane's current position. This is found by iterating through the possible values of θ_1 from 0 to θ_c in Equation 4, and stopping once the desired value of X is achieved. Finding the incident angle corresponding to the direct sound path at each point in the plane's trajectory will be important to solving the geometry proposed by Urick, discussed in the next section.

B. Two Iso-Speed Media Analysis Approach

The 1972 work by R. J. Urick has been presented as the basis for the formulation of this problem[1,2]. Reproduced in this solution's Figure 2 is the second figure from this work by Urick, which describes the geometry of the direct acoustic path from a moving plane source in the air medium to a submerged hydrophone[2]. Here, Urick proposes the placement of a "virtual source" in the water medium to completely model the real source in the air medium, such that the sound field at the hydrophone position from the real and virtual sources would be identical. Therefore, this problem can be reduced to one medium by tracing the trajectory of a virtual source in the water medium.

1. Virtual Source in Water Medium

Urick's proposed virtual source placement exists along a locus, which is modeled as being within a vertically-extended section of the water medium. It should be noted that Urick recommends a geometric solution to the description of this locus, as a mathematical representation is rigorous.

In this Figure 2, Urick proposes distance limits on the locus. The highest position on the locus above the ocean's surface and the largest horizontal displacement along the locus are limited by $(\frac{c1}{c2})h$. This locus describes the possible positions for placement of the virtual source, and thus the highest the virtual source will reach above the ocean's surface and the farthest horizontal displacement the virtual source will achieve relative to the plane's position are both bounded. As shown, this limit is proportional to the speeds of sound in the two media and the height of the plane, making the locus plane-height dependent.

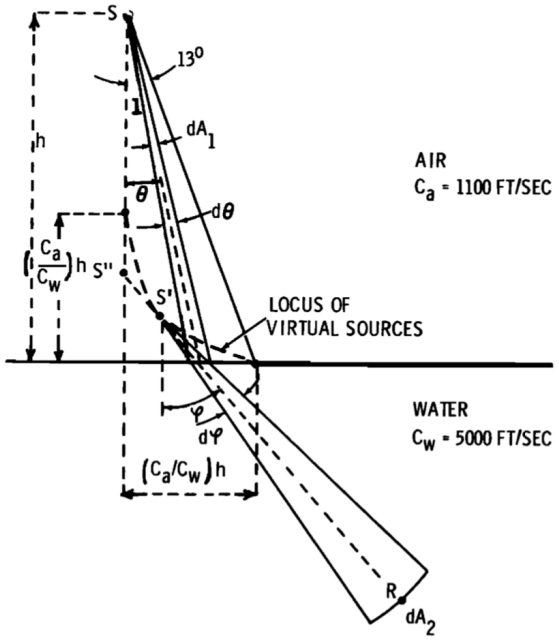


FIG. 2. Replication of Figure 2 from Urick 1972, showing the locus of virtual source locations based on plane position[2]

2. LUT Implementation of Virtual Source Position

For this solution, I will propose a computational approach to the generation of this locus. The first step to define this locus is to sweep through all possible incident angles from 0 to θ_c, and find the refracted angle into the water medium using Snell's Law. A refracted ray is then defined by this refraction angle with its tail at the intersection point of the incident acoustic ray and the air-water boundary. Then, each refracted ray is inverted back into the air medium. Figure 3 shows the shape created by the superposition of all of these rays when this is done computationally using MATLAB.

Each of these rays will be dominant at some point along the locus's curve, therefore, the overall shape of the locus can be found by selecting the maximal value among all the rays for each horizontal position. This also means that each horizontal and vertical position along the locus can be mapped to the corresponding dominant refracted ray, which can in turn be mapped to the incident acoustic ray from the plane source. Therefore, by knowing the incident angle corresponding to the direct sound path and the height of the plane, the relative horizontal and vertical coordinates based on the plane's position can be found by using a look-up table (LUT). The proposed nature of this row-vector LUT is as follows:

$$\text{LUT} = \begin{bmatrix} \text{Horizontal Position Offsets;} \\ \text{Vertical Positions;} \\ \text{Incident Angle to Oceans;} \\ \text{Refracted Angle into Oceans;} \\ \text{Ray Number} \end{bmatrix}$$

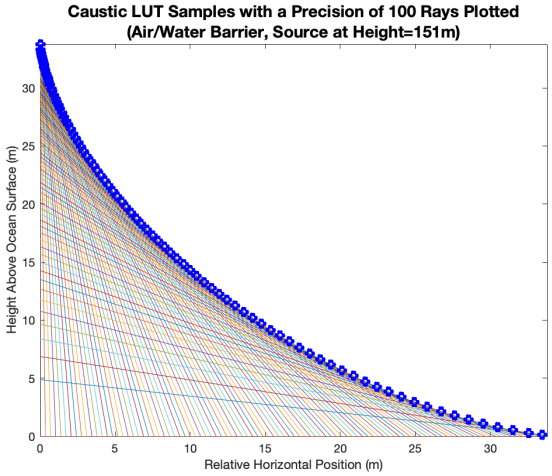


FIG. 3. Superimposed refracted rays showing the shape of the locus with a plane height of 151 meters

Thus, by knowing either the incident or refracted angle corresponding to the direct sound path, the virtual source can be placed relative to the current position of the plane's trajectory.

The final row in the LUT, the ray number, implicates the "precision" of the LUT. The incident angle range from 0 to θ_c is broken up into a certain number of equally spaced angles, each of which yields a refracted ray. The ray number in the LUT is the index in the incident angle range that yielded the dominant ray at that horizontal position. Figure 3 shows the dominant ray at each horizontal location denoted by a blue marker. This figure uses a precision of 100 rays, however the LUTs used for generating the results of this challenge problem will use two orders of magnitude more rays, yielding a more precise virtual source placement. `(An animated version of this LUT generation and ray tracing is available at TreDiPassio.com/ASA2019)`

3. Locus as a Caustic Curve

A final note on this locus of possible virtual source positions is that the shape describing this locus is caustic in nature. Caustic curves are prevalent in many fields where ray-tracing is of importance, such as optics. Caustics are very closely related to exponential functions. Therefore, as the trajectory of the virtual source will be determined by traversing the locus, it is expected that the derivative of the virtual sources position will closely resemble a scaled version of this caustic shape. This will be explored in Section II E.

C. The Doppler Effect

Using the virtual source placement described in the previous sections, for a given plane trajectory, a corresponding virtual source trajectory can be mapped. By knowing the positions through time of this virtual source, a derivative can be taken to determine the virtual source's velocity. Therefore, it is convenient to relate the velocity vector of a source to the instantaneous frequency.

1. Doppler Effect With On-Axis Source

Equation 5 relates wavelength to frequency in a medium with known sound speed[4]:

$$f = \frac{c}{\lambda} \tag{5}$$

where f is the frequency of radiation and λ is the wavelength of the propagation in that medium. If the source is moving on-axis with the receiver, the perceived wavelength at the receiver position will expand or compress based on the velocity v_s of the source[4]. If the source is moving toward the receiver, by convention $v_s = +|v_s|$, while $v_s = -|v_s|$ is used for a source moving away. Therefore, the following relationship is derived for moving source:

$$\lambda_{new} = \frac{c - v_s}{f} \tag{6a}$$

$$f_{new} = \frac{c}{\lambda_{new}} \tag{6b}$$

$$f_{new} = \frac{c}{\left(\frac{c-v_s}{f}\right)} \tag{6c}$$

$$f_{new} = \frac{c}{c - v_s} f \tag{6d}$$

Therefore, by Equation 6d there is a relationship for the change in perceived frequency at a stationary receiver location from a source moving with on-axis velocity v_s.

2. Doppler Effect in 2-D

In order to extend Equation 6d to two dimensions with a source moving at an angle θ relative to the stationary receiver, the compression and expansion of the propagating wavelength at the receiver position isn't directly along the axis of the source's motion[4]. Therefore the source's velocity vector can be decomposed to its components, and the overall source velocity used can be multiplied by the resultant factor of $cos\theta$, yielding Equation 7.

$$f_{new} = \frac{c}{c - |v_s| \cos\theta} f \tag{7}$$

As a sanity check, in the limit that the motion returns on-axis with the receiver, $\theta \to 0$, thus $\cos\theta \to 1$ and Equation 6d is re-obtained. Because the trajectory of the virtual source will follow an irregular path that is not on-axis with the receiver, Equation 7 will be used to obtain the instantaneous frequency perceived at the hydrophone position.

II. MODEL IMPLEMENTATION

Figure 4 shows the algorithm by which a model for the instantaneous frequency perceived at the hydrophone position is generated. Each of the processes involved will be described in this section.

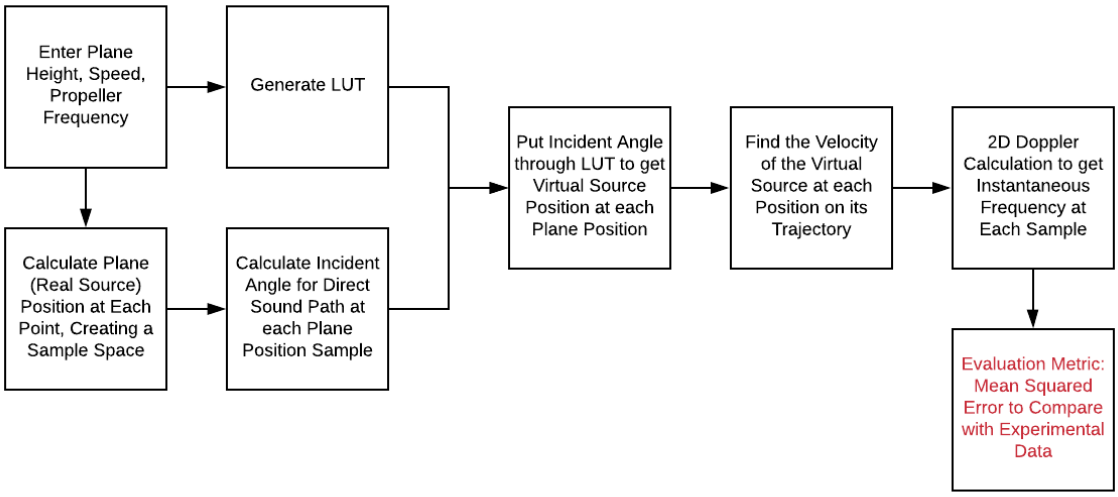


FIG. 4. Algorithm flowchart for creating model that predicts instantaneous frequency at the hydrophone position

A. Entering Plane Properties and Generating LUT

The problem statement provides three tasks to complete. In the first two, the plane's height, propeller frequency, and speed are provided with a given hydrophone depth. Therefore, the models generated through this algorithm for Tasks 1 and 2 will be created using these given parameters such that they can compared to the experimental data.

As described in Section I B 1, the caustic shape described in the LUT is plane-height dependent, therefore for each unique plane height, a new LUT will be generated using the aforementioned method from Section I B 2. For the rest of Section II, parameters from Task 1 will be used to demonstrate the output of the algorithm, and therefore a plane height of 151 meters, hydrophone depth of 20 meters, propeller frequency of 68 Hz, and a plane speed of 123 $^{m}/_{s}$ will be used.

B. Sampling the Plane's Trajectory

Sampling the plane's trajectory is straightforward yet necessary for the computational approach to this solution. Provided the hydrophone is located at some depth d, a two-dimensional cross section of the ocean and air above it, as portrayed in Figure 1, can be specified with a coordinate system such that the hydrophone is at point $(0,d)$. The plane flies directly over the hydrophone, beginning at some horizontal distance $-X$ from the position directly above the hydrophone, and ending at some horizontal distance $+X$. Therefore, for plane height h, within this coordinate system the plane flies from $(-X,h)$ to $(+X,h)$. These positions can then be sampled every hundredth of a meter to get an array of the plane's positions throughout its trajectory. Because the plane speed is known, abiding by the convention that the plane flies directly overhead the hydrophone at time of 0 seconds, each sampled position in the trajectory can timestamped.

C. Calculating Incident Angle at each Sample

For each position in the array containing the plane's trajectory, Equation 4 can be used to find the incident angle corresponding to the direct sound path. Figure 5 shows a plot of these angles using Task 1 parameters. This figure demonstrates that there are two distinct regions in this situation, an asymptotic region where the plane is far enough away in either direction that the angle change is negligible, and the region where the plane is flying directly overhead in which the incident angle transitions from θ_c to $-\theta_c$.

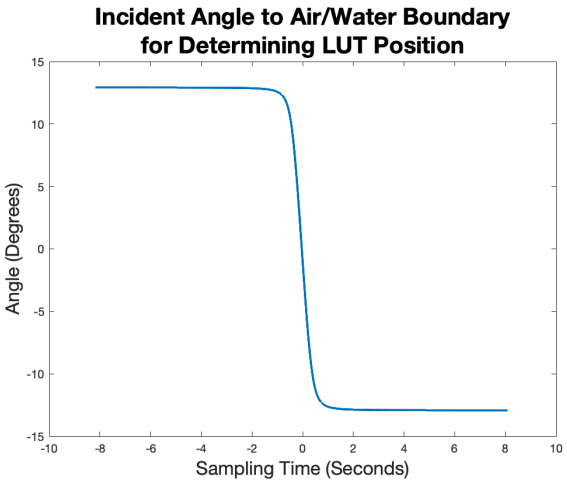


FIG. 5. Incident angles for the direct sound path

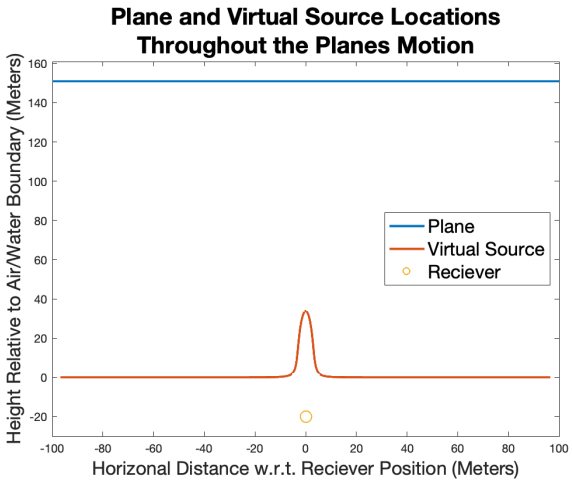


FIG. 6. Plane and virtual source trajectories relative to the hydrophone position

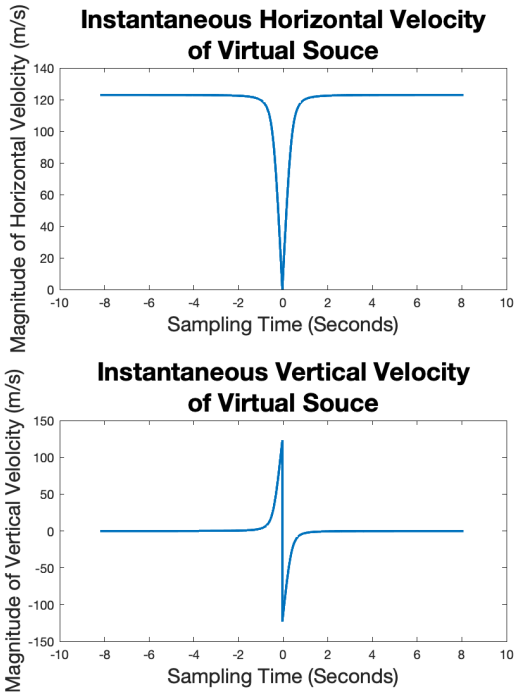
D. Determining the Virtual Source's Trajectory

From Figure 5, the incident angle for the direct sound path from plane to hydrophone at each timestamp is known. From the LUT described in Section I B 2, the vertical position and relative horizontal offset from the plane's position of the virtual source can be obtained directly from this incident angle array. Therefore, the virtual source's trajectory can be obtained with this data, and plotted alongside the plane's and the hydrophone's positions. This is seen in Figure 6. `(An animated version of this figure showing the plane and virtual source flying in time is available at TreDiPassio.com/ASA2019)`

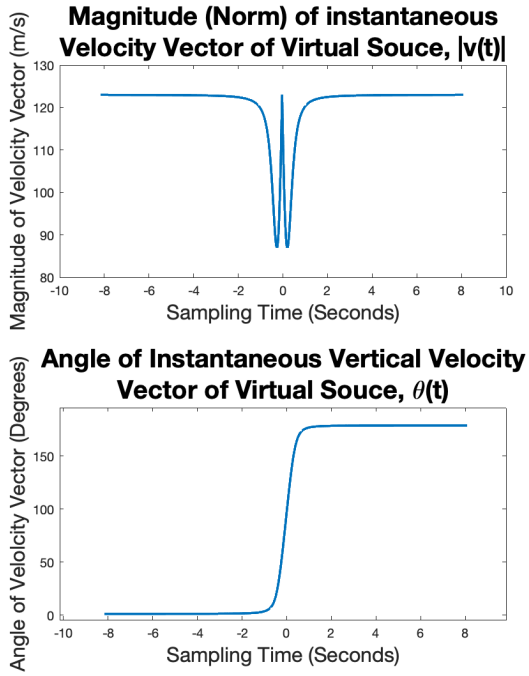
The trajectory of the virtual source looks as would be expected. While the plane is far away in either direction, there is little vertical motion in the virtual source's trajectory as there is little change in the incident angle. As the plane flies overhead, the LUT is quickly traversed, and the caustic nature of the LUT is mirrored in the virtual source's trajectory.

E. Finding the Virtual Source's Velocity

Since the virtual source's position is known for each point in time, the derivative of this trajectory will yield velocity information. As this is a discretely sampled trajectory, difference equations in the horizontal and vertical directions will compute this differentiation, yielding component velocity vectors which are shown in Figure 7(a). Intuitively, both of these components should be caustic in nature as explained in Section I B 3. The horizontal component should go to zero as the plane flies overhead and should converge to the magnitude of the plane's speed in the asymptotic regions. The vertical component should go to zero in the asymptotic regions and should saturate to the plane's speed when directly above the hydrophone.



(a)



(b)

FIG. 7. (a) Horizontal and vertical components of the instantaneous velocity vector, (b) Norm and angle of instantaneous velocity vector

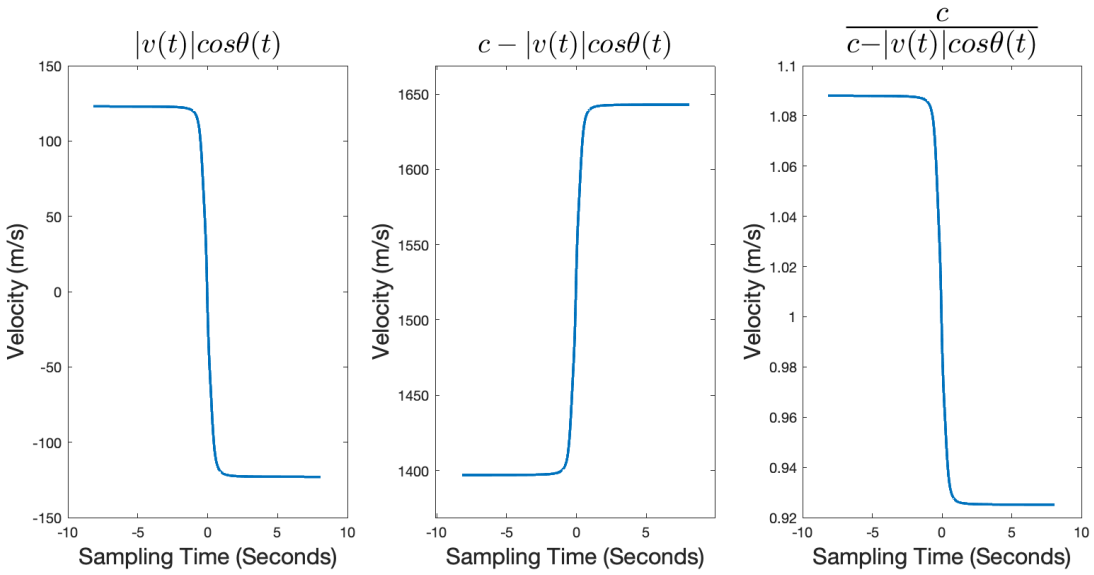


FIG. 8. Building the coefficient from Equation 7

As the virtual source begins its "descent" directly after the overhead sample, the vertical velocity component will change sign. These expectations are demonstrated by the plot of these components.

By isolating these components, the norm of the vector can be taken at each point to determine the magnitude of the velocity, and the dot product can be used to find the angle of the velocity. This is shown in Figure 7(b). As expected, the angle of the instantaneous velocity vector transitions from 0° during the plane's approach to 180° during the transition region, and remains stable in the asymptotic regions.

F. Instantaneous Frequency Calculation

The relationship relating the velocity vector to the instantaneous frequency is detailed by Equation 7, whereby the Doppler-shifted frequency is seen as the true source frequency scaled by the coefficient $\frac{c}{c-v_s \cos\theta}$. This coefficient is evaluated using the norm and angle of the instantaneous velocity vector presented in Figure 7(b). It is built element by element in Figure 8 until the final shape of the coefficient through time is observed. This curve is then multiplied by the true propeller frequency to obtain the final predicted instantaneous frequency curve shown in Figure 9. Here, it should be noted that the original propeller frequency of 68 Hz is observed when the plane is directly overhead the hydrophone, by convention at a timestamp denoted as 0 seconds. As expected, asymptotic values of the instantaneous frequency are observed when the plane is far away, and a rapidly changing transition region is seen as the plane flies overhead.

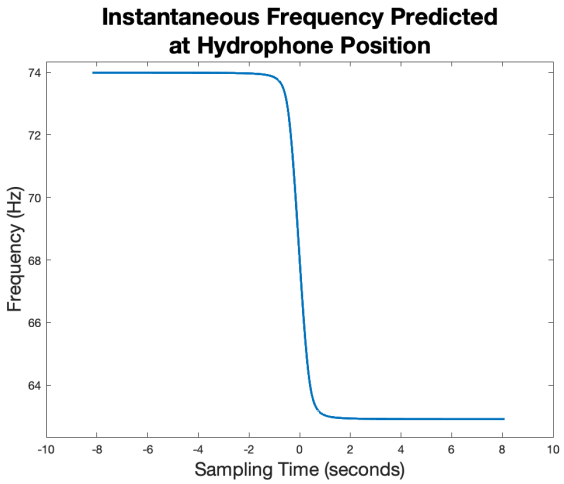


FIG. 9. Predicted instantaneous frequency at the hydrophone position with Task 1 plane parameters

III. COMPARISON TO EXPERIMENTAL DATA

The previous section demonstrated the use of an algorithm to develop a model for the predicted instantaneous frequency based on the plane's speed, height, propeller frequency, and the hydrophone's depth. The first two tasks in this challenge problem ask for this model to be evaluated with given experimental data. This section will show these comparisons using a mean squared error approach.

A. Mean Squared Error

The evaluation metric I will use to describe the closeness of fit of the predictive model and the experimental data is mean squared error (MSE)[5]. MSE describes the average squared deviation between the model and experimental values. The following equation shows this method[5]:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - Y_i^{'})^2 \tag{8}$$

where n is the number of points compared, Y_i is the experimental value at the ith compared point, and $Y_i^{'}$ is the predicted value at that index i. This evaluation will yield an answer in the units provided squared. Therefore, a secondary comparison, the root mean squared error (RMSE), which is the square root of the MSE, can be used to describe the deviation in the same units as the samples[5].

Because both the given experimental values and the predicted instantaneous frequency values from the generated model are timestamped, the model can sampled at the same timestamps at which the experimental data is sampled, and these two arrays can be passed to a MATLAB function to evaluate Equation 8 using synchronized data.

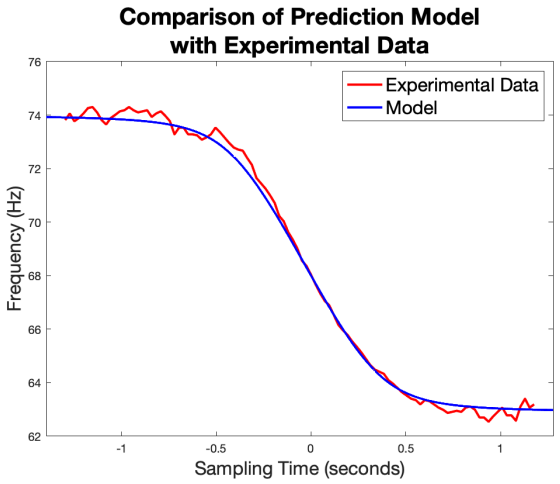
B. Task 1

Task 1 involves the comparison of the generated model to a given curve that charts observed frequency values through time at a real hydrophone position as a plane flies overhead. The parameters given in Task 1 are a plane height of 151 meters, hydrophone depth of 20 meters, propeller frequency of 68 Hz, and a plane speed of 123 $^{m}/_{s}$.

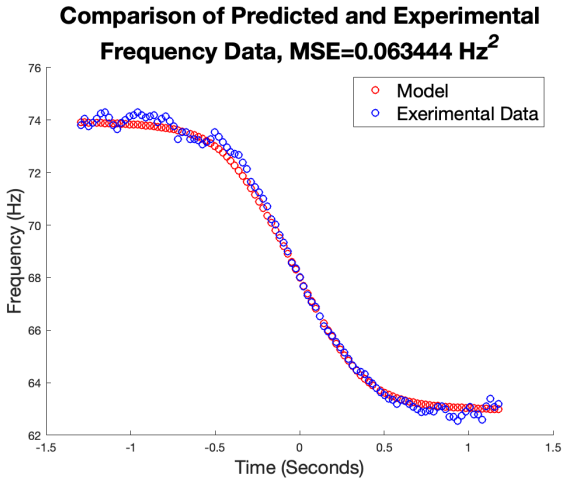
The first step is to execute the algorithm shown in Figure 4 using these parameters. Because these parameters were used in the previous section to demonstrate the algorithm's execution, Figure 9 shows the final model created for this Task. The given experimental data and the generated model are plotted on the same time axis in Figure 10(a).

Anecdotally, it appears that the model's predicted asymptotic regions and transition region are good fits to the experimental data. This is quantified in Figure 10(b) when the MSE is evaluated at the synchronized timestamps as described in Section III A.

The total MSE between the predictive model and the experimental data is 0.06344 Hz^2. This means that the RMSE is 0.252 Hz. The highest reported frequency from the experimental data is 74.3 Hz, and the minimum reported value is 62.54 Hz, giving a total frequency range spanning 11.76 Hz. The average deviation of the model to the experimental data, reported by the RMSE, is thus 2.13% of the total experimental frequency range. This value shows the predictive model as an accurate estimator of the actual instantaneous frequency seen at the hydrophone position.



(a)



(b)

FIG. 10. (a) Comparison of the given experimental instantaneous frequency curve with the curve generated by the predictive model (b) MSE comparison of experimental and model data using synchronized data samples

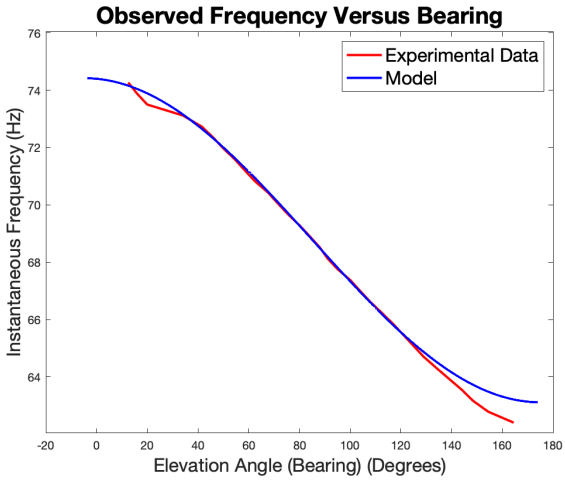
C. Task 2

For Task 2, the experimental data is given in the form of instantaneous frequency versus apparent bearing of the incoming acoustic ray in the direct sound path. The definition of bearing for this problem is defined in the problem statement as the complement to the refracted angle in the water medium. The given parameters for the second task are slightly different than those given in the first task, with the propeller frequency at 68.3 Hz and a plane speed of 125 $^{m}/_{s}$. Because the instantaneous frequency will be tested versus apparent bearing, the plane height and the hydrophone depth shouldn't impact the Task 2 results, as they should only impact the width of the transition region by changing the times at which in-

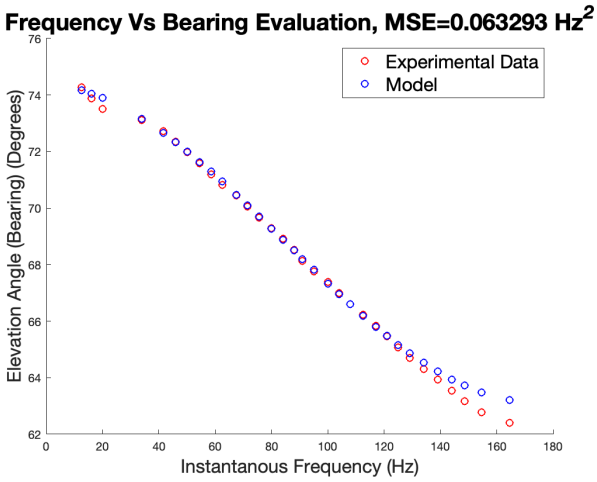
cident angles transition. Thus these parameters are kept the same as in Task 1 for simplicity.

Firstly, the model with these parameters is generated with the algorithm described in Figure 4. Because the refracted angle at each point is a row in the LUT, the refracted angle and its compliment (bearing) can be logged through time and paired one-to-one with the instantaneous frequency measurements in the model. This frequency versus bearing relationship is plotted alongside the given experimental data in Figure 11(a). As before, anecdotally this looks to be a good fit, which can again be quantified using the mean squared error.

In similar fashion to the process described in Section III A, the instantaneous frequency measurements are sampled at the same bearings which are reported with the experimental data, and plotted in Figure 11(b). This allows for the calculation of the MSE at the appropriate sample locations, which is reported to be 0.063293 Hz^2. This corresponds to an RMSE of 0.251 Hz, very consistent with the results from the previous task. The frequencies in the given experimental data for this task range from 62.4 Hz to 74.27 Hz, yielding a total frequency range of 11.87 Hz. Therefore, the average deviation of the model to the experimental data, again reported by the RMSE, is 2.12% of the total experimental frequency range, showing again that the predictive model is an accurate estimator of the actual instantaneous frequency seen at given bearing values.



(a)



(b)

FIG. 11. (a) Comparison of given experimental frequency vs bearing curve with the curve generated by the predictive model (b) MSE comparison of experimental and model data at synchronized bearing samples

IV. TASK 3

The nature of the third and final task of this challenge problem is different than that of the previous two. The third task provides unlabeled raw audio data recorded from the hydrophone and asks for the extraction a recorded plane's height, speed, and propeller frequency. Figure 12 shows a flowchart of the approach used to solve this problem utilizing the model generation algorithm previously described. This flowchart will be explained in the following sections.

A. Extracting Data from Provided Audio

The file provided for this task contains 120 seconds of audio recorded by the hydrophone, sampled at 44100 Hz. In order to extract instantaneous frequencies from the provided file, a short time Fourier transform (STFT) is utilized[6]. The audio is split into frames with a length of 8192 samples per frame, windowed with an 8192 point Hamming window. Each frame is converted to the frequency domain using the fast Fourier transform (FFT) algorithm run on the frame data[6]. 75% overlap is used between the frames, and zero padding is added such that the frequency resolution (bin size) of the FFT in each frame is approximately 0.1 Hz. Consistent with the Nyquist Theorem, the maximal frequency present in the given audio is half the sampling frequency, or 22050 Hz. Therefore, the full-scale 22050 Hz will be represented in the number of frequency bins equal to half of the FFT length due to the redundancy added when computing the FFT. The following equation thus shows how the desired resolution is achieved[6]:

$$\frac{f_s}{2} = \frac{\frac{1}{2}\text{FFTLEN}}{\text{Resolution}} \rightarrow \text{Resolution} = \frac{f_s}{\text{FFTLEN}} \quad (9)$$

Solving Equation 9 for the desired resolution of around 0.1 Hz in each frame shows that 53-times zero padding should be used.

The maximum frequency bin of each frame is used to estimate the fundamental frequency present in that frame. Figure 13 shows these estimates, as well as the root-mean-square (RMS) amplitude computed in each frame. The majority of the audio solely contains underwater noise. However, at around 54 seconds, a few

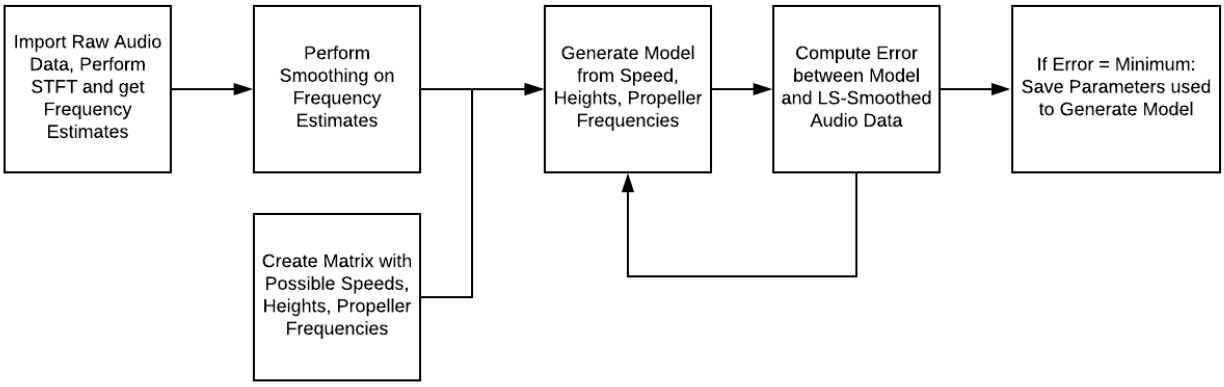


FIG. 12. Algorithm flowchart for determining model parameters that minimize MSE when compared to experimental data

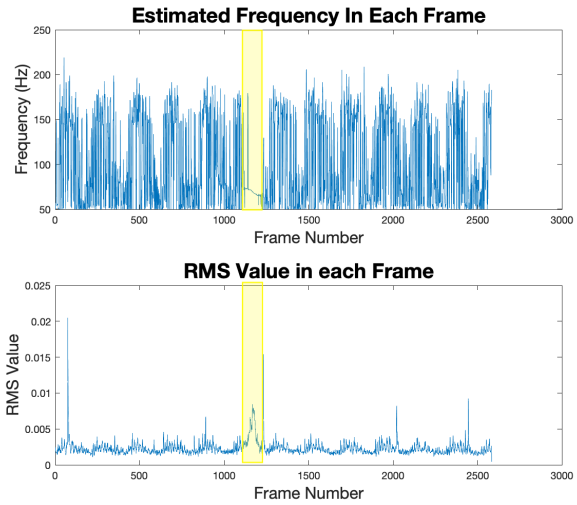


FIG. 13. Estimated fundamental frequency and RMS amplitude in each frame of the given audio file. The frames where the plane is flying overhead are highlighted

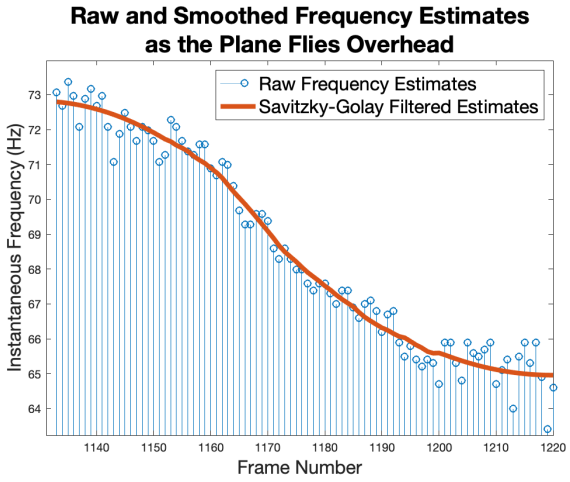


FIG. 14. Raw and smoothed frequency track at the hydrophone position extracted from the given audio data of the plane flying overhead

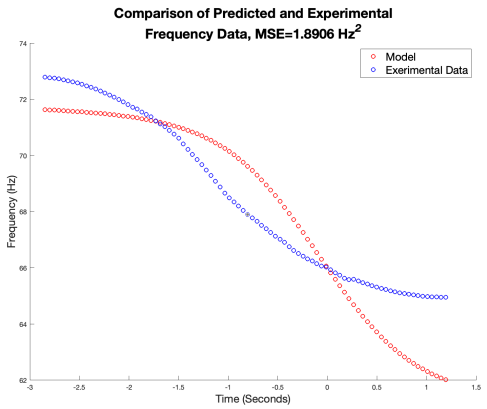
seconds of a plane flying overhead can be heard when listening to the file. The presence of the plane is shown in the highlighted section of Figure 13. There is a spike in the RMS amplitude above the noise floor, paired with a more defined instantaneous frequency track. Figure 14 shows only the frames where the plane flying overhead is audible commiserate to the RMS spike in Figure 13.

The estimates from each frame in the region with audible plane motion are smoothed to create an instantaneous frequency track. This smoothing is done with a Savitzky-Golay filter[7,8]. This type of filter is used to smooth sets of data points in the the time domain using a least-squares fitting approach[7].

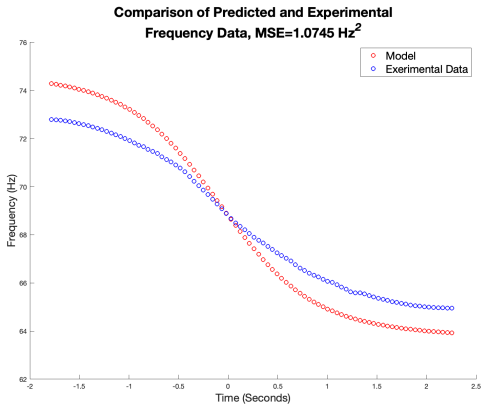
B. Generating Model Information Matrix

The general approach will be to find the plane height, speed, and propeller frequency that, when a model is implemented using those parameters, the predicted instantaneous frequency track most closely resembles the smoothed frequency estimates in Figure 14. That is to say, the model minimizes the MSE among all models tested. Because the possible speeds of the aircraft, propeller frequencies, and the heights of the plane come from a relatively limited set, a brute force method is proposed whereby each possible combination is tested. Note that the hydrophone depth is 90 meters in all generated models for this task.

While the overall computational time could be reduced by using any of various optimization methods, the



(a)



(b)

FIG. 15. Comparison of experimental data with models generated by poor-fitting parameters, in which (a) all three parameters are shown to be poorly estimated leading to poor transition width, asymptotic behavior, and synchronization (b) height and propeller frequency are close to correct, shown via appropriate transition width and synchronization, however poorly-matching asymptotic behavior shows plane speed is incorrectly estimated

limited nature of the set of possible parameters will allow local minima traps to be avoided entirely by carrying out complete testing. Plane speeds between 50 m⁄s and 200 m⁄s were tested, incremented every 1 m⁄s. Heights between 50 and 200 meters were tested, incremented every meter. Propeller frequencies between 64 Hz and 73 Hz were tested, incremented every 0.1 Hz. After the optimal parameters from this set were extracted, a second smaller set of parameters surrounding the initial estimates were tested in which the plane speed is incremented every 0.1 m⁄s and the plane height is incremented every 0.1 meters to increase the overall precision of the parameter search.

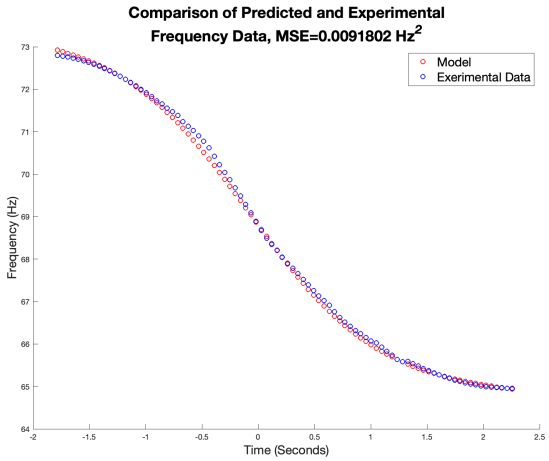


FIG. 16. Comparison of the experimental data with the model generated by the extracted plane parameters that create a frequency track which minimizes the MSE. Here, plane height is 153.7 meters, plane speed is 100.9 m⁄s, and propeller frequency is 68.8 Hz. These parameters will be reported in Section IV D as the final answers to Task 3

C. Model Generation and MSE Evaluation

For each possible trio of plane height, speed, and propeller frequency from the sets described in the previous section, a model is generated using the algorithm from Figure 4. Because the sampling frequency of the provided audio is known, the amount of overlap used when segmenting the audio into frames, as described in Section IV A, can be used to create relative timestamps for each frame. Then, these timestamps can be synchronized with the timestamps from the generated models using the convention that the propeller frequency is observed at a time of 0 seconds, as described in Section III A. The MSE of each model's predicted frequency track when compared to the smoothed frequency track shown in Figure 14 is recorded. After all parameters are tested in the manner described in the previous section, the parameters that resulted in the minimized MSE are in theory the parameters closest to those that generated the raw audio given, and therefore those are the parameters that will be reported in the next section as the answer to Task 3.

D. Final Extracted Parameters

Using the process detailed in the previous section, the parameters that generate the model that most closely resembles the smoothed frequency track observed in Figure 14 can be extracted via minimization of MSE. Firstly, a selection of poor model fits are shown in Figure 15 to demonstrate why this optimization is necessary. In 15(a), all three parameters are estimated incorrectly. The speed and the height of the model are incorrect, leading to improper asymptotic behavior and a wider transition respectively. The propeller frequency is also improperly es-

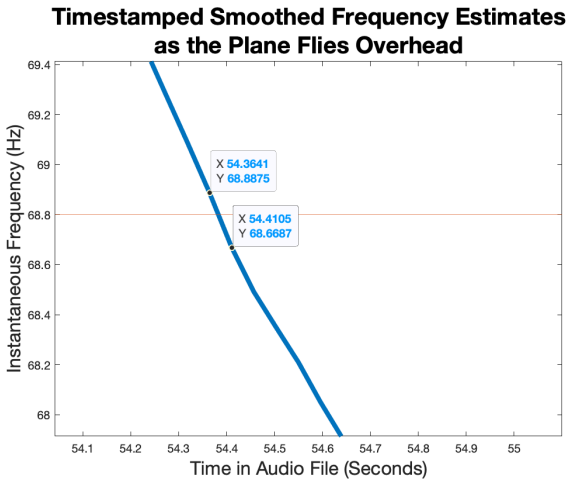


FIG. 17. Zoomed in and timestamped frequency estimates with extracted propeller frequency "sandwich points" labeled

timated, leading to poor synchronization in the samples. 15(b) shows improper asymptotic behavior, but appropriate synchronization and transition width. Therefore, the height and propeller frequency are more appropriately estimated, but the speed parameter is incorrect.

When the algorithm shown in Figure 12 is completely carried out, the minimized MSE is reportedly achieved with the following parameters: **plane height of 153.7 meters, plane speed of 100.9 m/s, and propeller frequency of 68.8 Hz**. Figure 16 shows the comparison of the model generated with these parameters to the estimated frequency track from Figure 14. Anecdotally, they are a good match. This can be quantified by the MSE, reported to be 0.0091802 Hz^2. This corresponds to an RMSE of 0.096 Hz. The experimental frequency values from Figure 14 range from 63.40 Hz to 73.37 Hz, for a total range of 9.97 Hz. Therefore, the average deviation of the model to the experimental estimates, once again reported by the RMSE, is 0.95% of the total frequency range. Therefore, these parameters are extracted and reported with a high confidence rating via the algorithm shown in Figure 12.

E. Sanity Check

Recall that, in the limit, the one-dimensional Doppler Equation 6d relates the true propeller frequency to the frequencies observed in the asymptotic region of the instantaneous frequency track. Therefore, the following expressions for the inbound asymptotic frequency, $f(-\infty)$, and the outbound asymptotic frequency, $f(+\infty)$ can be shown following from Equation 6d and the aforementioned conventions as follows:

$$f(+\infty) = \frac{c}{c+|v|} f_0. \tag{10a}$$

$$f(-\infty) = \frac{c}{c-|v|} f_0 \tag{10b}$$

From the experimental data, $f(-\infty)$ is expected to be 73.37 Hz and $f(+\infty)$ is expected be 63.40 Hz. Following Equations 10a and 10b, the following sanity check can be performed using the extracted plane parameters:

$$f(+\infty) = \frac{1520^{m}/_{s} * 68.8\text{Hz}}{1520^{m}/_{s} + 100.9^{m}/_{s}} = 64.51\text{Hz} \tag{11a}$$

$$f(-\infty) = \frac{1520^{m}/_{s} * 68.8\text{Hz}}{1520^{m}/_{s} - 100.9^{m}/_{s}} = 73.69\text{Hz} \tag{11b}$$

Thus, the extracted speed and propeller frequency parameters can be used to evaluate the asymptotic behavior of the experimental frequency track analytically. The percent error in the calculated $f(+\infty)$ from Equation 11a is 1.75 %. The percent error in the calculated $f(-\infty)$ from Equation 11b is 0.44 %. Therefore, this sanity check whereby the calculated asymptotic behavior matches the experimental frequency track within 2% error helps assure confidence in the extracted speed and propeller frequency estimates.

F. Timestamp where Plane is Overhead

As mentioned previously, by knowing the sampling frequency of the provided audio data and knowing the frame size and overlap points used to segment this audio into frames, each of these frames can be timestamped. A zoomed-in snippet of where the smoothed experimental frequency curve crosses the extracted propeller frequency of 68.8 Hz is plotted in time (rather than in frame numbers) in Figure 17. The so-called "sandwich points" are labeled, such that linear interpolation can be done to estimate where exactly this 68.8 Hz frequency in achieved. From the labeled data points, the linear interpolated position of what would be the 68.8 Hz sample is found to be 54.383 seconds relative to the beginning of the recording. **Therefore, the audio from the plane directly overhead is calculated to be heard at 54.383 seconds in the given recording**.

However, sound travels at a finite speed. Therefore, while the 68.8 Hz sample might be `heard` at 54.383 seconds, it wasn't emitted from the plane's propellers at that same timestamp. Instead, the speed of sound and distance of travel in the two media allow the true timestamp at which the plane flies overhead to be calculated. The following equation shows this:

$$\text{Propagation Time} = \frac{h}{c_1} + \frac{d}{c_2} \tag{12}$$

Thus for the speeds of sound in air and water, the extracted height parameter of 153.7 meters, and the hydrophone depth of 90 meters, Equation 12 yields a propagation time of 0.511 seconds from plane to receiver when directly overhead. Therefore, the true timestamp that the plane was directly above the hydrophone is 0.511 seconds before the sample at 54.383 seconds. **Therefore, the plane was directly above the hydrophone at 53.872 seconds in the recording.** Answers from all tasks will be compiled in the next section.

V. FINAL REPORTED ANSWERS

For Task 1, Figure 10 shows the degree to which the experimental and model-generated instantaneous frequency curves match each other, with a reported MSE of 0.06344 Hz^2. This translates to an RMSE of 0.252 Hz, which is 2.13% of the total experimental frequency range from the given data.

For Task 2, Figure 11 shows the degree to which the experimental and model-generated instantaneous frequency versus bearing curves match each other, with a reported MSE of 0.063293 Hz^2. This translates to an RMSE of 0.251 Hz, which is 2.12% of the total experimental frequency range from the given data. Task 1 and Task 2 show that the generated model is consistent and accurate when predicting the instantaneous frequency curve at the hydrophone position for the given plane parameters.

For Task 3, the algorithm shown in Figure 12 was carried out, and the plane parameters that generated the model which minimized the MSE relative to the experimental data were as follows: **plane height of 153.7 meters, plane speed of 100.9 $^m/_s$, and propeller frequency of 68.8 Hz**. Figure 16 shows the degree to which this model and the experimental frequency curve match each other, with a reported MSE of 0.0091802 Hz^2. This translates to an RMSE of 0.096 Hz, which is is 0.95% of the total experimental frequency range from the given audio data. The timestamp where the true propeller frequency is observed in the audio recording was calculated to be at **54.383 seconds**. By calculating the acoustic propagation time from the plane's propellers to the hydrophone position, the true time the plane was directly above the hydrophone was found to be **53.872 seconds.**

With a brief sanity check completed in Section IV E, these parameters represent my best and final prediction for the true plane parameters for the plane recorded in the given audio file.

[1] B. G. Ferguson, R. L. Culver, and K. L. Gemba, "International Student Challenge Problem in Acoustic Signal Processing 2019," Acoustics Today **15**(1), 71–73.

[2] R. J. Urick, "Noise Signature of an Aircraft in Level Flight over a Hydrophone in the Sea," The Journal of the Acoustical Society of America **52**(3B), 993–999 (1972) asa.scitation.org/doi/abs/10.1121/1.1913206 doi: 10.1121/1.1913206.

[3] T. Rossing and N. Fletcher, "Sound waves in air," in *Vibrations and Sound*, 2 ed. (Springer-Verlag, New York, 2004), Chap. 6, pp. 139–156.

[4] M. Schwartz, "Lecture 21: The Doppler Effect," users.physics.harvard.edu/~schwartz/15cFiles/Lecture21-Doppler.pdf.

[5] G. Drakos, "How to select the Right Evaluation Metric for Machine Learning Models: Part 1 Regression Metrics," Medium (2018).

[6] D. Arfib, F. Keiler, U. Zölzer, V. Verfaille, and J. Bonada, "Time-frequency processing," in *DAFX: Digital Audio Effects*, edited by U. Zölzer, 2 ed. (John Wiley and Sons, Inc., 2011), Chap. 7, pp. 219–277.

[7] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing* (Cambridge University Press, New York, NY, USA, 1988), pp. 650–655.

[8] "Smooth noisy data - MATLAB smoothdata," www.mathworks.com/help/matlab/ref/smoothdata.html.