

COVER PAGE

Document downloaded by @DAEL

Fri Nov 27 16:28:44 2026
For personal use

When automatic English translation is provided, only the original document is authentic.
The EAA cannot be held responsible of any translation error

Bibliographical reference

Passive Acoustic Method for Tracking Moving Sound Sources, M. Genescà, J. Romeu, T. Pàmies and A. Balastegui, Acta Acustica **vol. 97** (Number 1), 2011, pp. 34-43

DOI
https://doi.org/10.3813/AAA.918384

ACTA ACUSTICA UNITED WITH ACUSTICA
Vol. 97 (2011) 34 – 43

DOI 10.3813/AAA.918384

Passive Acoustic Method for Tracking Moving Sound Sources

M. Genescà, J. Romeu, T. Pàmies, A. Balastegui
Laboratory of Acoustics and Mechanical Engineering (LEAM), Technical University of Catalonia, C/ Colom 11, 08222 Terrassa, Spain. meritxell.genesca@upc.edu

Summary
This paper presents an acoustic method to obtain the instantaneous velocity and position of a moving sound source following an unknown 3D path provided that the position of the source is known at a random time instant. The method involves a mesh of at least 7 acoustic receivers distributed in the acoustic influence area of the source, and assumes that the background noise is uncorrelated between them. It is necessary to know the position of the source at a random instant in order to compensate for the propagation time delay between the signals of the receivers. The method focuses on the relative Doppler Effect between a reference microphone and the rest to obtain the velocity and position at every time instant. The performance of the method has been simulated and experimentally tested in the case of a moving sound source travelling in a linear path, and the location errors obtained in both cases are acceptable.

PACS no. 43.28.Tc

1. Introduction

The object of this paper is to present a passive acoustic method to obtain the instantaneous velocity and position of a moving sound source following an unknown 3D path provided that the position of the source is known at a random time instant. Therefore, this method would be useful, for instance, in cases where the departure position of the source is known and the rest of the path is unknown.

An example application is the tracking of aircraft near an airport, as the position of the aircraft is known at the time of takeoff. In the case of smaller airfields which deal primarily with light aircraft, and don't have a RADAR system, this method is an alternative for tracking the aircraft. In the case of major airports, this method can utilize spectral data from existing noise monitoring terminals to both track the aircraft and monitor noise levels. Therefore, the noise monitoring system becomes independent of non-acoustical data (such as RADAR data) that is often restricted or released with delay for security purposes. In both cases uncorrelated background noise in the receivers is present, and so is considered in the method.

The problem of tracking a moving sound source differs from that of locating a stationary (relative to the observer) sound source. In the first case, the signals involved are affected by the Doppler Effect. Therefore, standard methods such as Time Difference of Arrival [1] can't be applied to a mesh of microphones spread over a large area because the signal of each microphone would be affected differently by the Doppler Effect and so the performance of time difference estimation would decrease. The Time Difference of Arrival is, therefore, only applicable in particular cases where the Doppler dilation is small.

Target tracking is widely studied by active SONAR techniques where a reference signal is emitted into the ocean, the echo of this signal propagates back from the target, from the water surface and from the bottom of the sea. This signal is received by a spatially distributed sensor (hydrophones) array that separates the signal as a function of its azimuth and vertical angle of arrival (i.e the possible direction of the target seen from the array). The signal received from each direction is compared to the emitted signal using the Ambiguity Function from which the Doppler frequency dilation and the time delay between the emitted and the received signals are obtained. On one hand, the target range is obtained from the time delay, and along with the target direction given by the array, states the target position. On the other hand, the target range rate is obtained from the Doppler frequency dilation, so the relative movement from the target to the receiver is also found [2]

The present work is aimed at developing a tracking system for acoustic sources, so there is no need to emit a reference signal since the target is already an emitter itself. Systems that use the target's own noise to locate it are called passive systems. Some passive sonar systems use a spatially distributed sensor array to locate the source from the time delay existing between the received signal in every sensor pair [3, 4, 5]. In this case the microphones are placed close enough together so that the influence of the Doppler Effect is the same for all of them. Techniques based on the time delay estimation between pairs of sen-

Received 16 March 2010,
accepted 10 November 2010.

© S. Hirzel Verlag · EAA

sors have also been attempted for on-land applications in the case of narrow band, broad-band and impulsive sound sources [6, 7, 8, 9]. Also sound intensity 3D- measurements have been used to estimate the source direction relative to the receiver [10]. Other methods using directional microphones completely locate the source from the sound amplitude [11]. When multiple sources may be present at the same time the previous methods don't work as well, and often microphone array beamforming is used [12]. To obtain the position of the source either a 3D microphone array, or several 1D or 2D arrays is needed, involving a complex processing algorithm, and often, a large number of microphones. Other methods are based on the measurement of the changes in the received frequency caused by Doppler Effect. Systems with just one microphone are used to obtain the minimum distance between the source and the receiver [13]. In systems that use a mesh of microphones, the Doppler frequency dilation affects differently each receiver and this fact can be used to locate an acoustic narrow band tone source [14]. In that case, the instantaneous frequency in each receiver can just be measured using short time Fourier transforms. The present method is also based in the Doppler frequency dilation but differs from the last since it can be also used for broadband noise sources by using the Wide Band Ambiguity function [15, 16] to measure the relative Doppler frequency dilation between each pair of microphones.

The present method involves a mesh of acoustic receivers distributed along the acoustic influence area of the source. When an acoustic signal propagates from a moving source to one of these static receivers, the signal is time delayed due to the propagation path, and time stretched due to the fact that the source is moving (Doppler Effect). Section 2 reviews these two effects. On one hand, if the initial position of the source is known, then the propagation delay can be calculated and removed from the received signal. On the other hand, the Doppler Effect causes frequency dilation in the received spectra; this dilation is related to the source position and speed. Section 3 of this paper explains how the relative frequency dilation can be calculated from the frequency spectra measured in every receiver, and Section 4 shows the geometric relations that allow the source to be located and its velocity calculated from the relative frequency dilation between a set of receivers. In section 5 the performance of the method is studied by simulation and experimental test, in the case of a sound source moving on a linear path at constant speed. Finally, section 6 contains the conclusions of this study.

2. Propagation from the source to the receiver

Given a random distribution of the receivers, let the vector $\boldsymbol{r}_n$ represent the position of the receiver n, and the vector $\boldsymbol{p}(t_0+t)$ represent the source position at a generic time instant t seconds after t_0. Then, the vector $\boldsymbol{q}_n(t_0+t)$ between the source and the receiver n is given by

$$\boldsymbol{q}(t_0+t) = \boldsymbol{r}_n - \boldsymbol{p}(t_0+t).$$

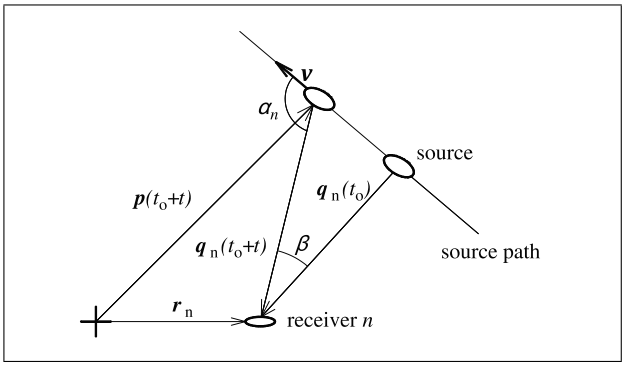

Figure 1. Vectors and angles involved in the sound propagation expressions from the source to the receiver.

In Figure 1, $\boldsymbol{v}$ represents the source velocity that is assumed to be constant within $0 \leq t \leq \Delta t$ provided that Δt is small enough. The angle α_n is the angle between v and $\boldsymbol{q}_n(t_0+t)$.

Taking the line between the receiver and the source as the moving reference frame, the speed of the source can be expressed as

$$\boldsymbol{v} = \dot{\boldsymbol{\beta}} \times \big(-\boldsymbol{q}(t_0+t)\big) + \dot{\boldsymbol{q}}_n(t_0+t), \tag{1}$$

where $\dot{\boldsymbol{\beta}}$ is the angular speed of the moving reference frame, $\dot{\boldsymbol{q}}_n(t_0+t)$ is the speed of the source relative to the moving reference frame, $\dot{\boldsymbol{q}}_n(t_0+t)$ is collinear with the line from the receiver to the source while $\dot{\boldsymbol{\beta}} \times (-\boldsymbol{q}(t_0+t))$ is perpendicular to $\boldsymbol{q}_n(t_0+t)$, then

$$\dot{\boldsymbol{q}}_n(t_0+t) = -\boldsymbol{v}\cos\alpha_n. \tag{2}$$

Using the same reference frame, the acceleration of the source is given by

$$\begin{aligned}\alpha \;=\; & -\dot{\boldsymbol{\beta}} \times \dot{\boldsymbol{\beta}} \times \big(-\boldsymbol{q}_n(t_0+t)\big) \\ & + \ddot{\boldsymbol{\beta}} \times \big(-\boldsymbol{q}_n(t_0+t)\big) + 2\dot{\boldsymbol{\beta}} \times \boldsymbol{v} + \ddot{\boldsymbol{q}}_n(t_0+t).\end{aligned} \tag{3}$$

If $\boldsymbol{v}$ is assumed to be constant within $0 \leq t \leq \Delta t$ and $\ddot{\boldsymbol{q}}_n(t_0+t)$ is collinear with $\boldsymbol{q}_n(t_0+t)$, then

$$\ddot{\boldsymbol{q}}_n(t_0+t) \;=\; \dot{\boldsymbol{\beta}}^2 \cdot \big(-\boldsymbol{q}_n(t_0+t)\big) - 2\boldsymbol{v}\dot{\boldsymbol{\beta}}\sin\alpha_n. \tag{4}$$

Moreover, if $\boldsymbol{v}\cdot\Delta t \ll \boldsymbol{q}_n(t_0)$ then $\beta \to 0$ and $\dot{\beta} \to 0$, thus $\ddot{\boldsymbol{q}}_n(t_0+t) = 0$ and the relative speed is constant. If the relative speed is constant, the distance between the receiver and the source is

$$\boldsymbol{q}_n(t_0+t) \;=\; \boldsymbol{q}_n(t_0) - \boldsymbol{v}\cos\alpha_n \cdot t. \tag{5}$$

If c is the speed of sound, the signal emitted s_e at a certain time instant t_0+t will arrive $\boldsymbol{q}_n(t_0+t)/c$ seconds later at the receiver n, therefore

$$y_n\left(t_0 + t + \frac{q_n(t_0+t)}{c}\right) = s_e(t_0+t). \tag{6}$$

The amplitude attenuation is omitted in equation (6), and in the rest of the paper, as it has no influence on the final

results because they depend on the shape of the signal, not on the amplitude.

Combining equations (5) and (6),

$$y_n\left(t_0 + t\frac{c - v\cos\alpha_n}{c} + \frac{q_n(t_0)}{c}\right) = s_e(t_0 + t). \quad (7)$$

Equation (7) shows the time at which the signal generated at $t_0 + t$ is received. However, it is more useful to know when a signal, received at time instant $t_0 + t'$, was generated. To obtain this information, the following change of variables is used,

$$t' = t\frac{c - v\cos\alpha_n}{c} + \frac{q_n(t_0)}{c}, \quad (8)$$

$$y_n(t_0 + t) = s_e\left(t_0 + \frac{c}{c - v\cos\alpha_n}\left(t' - \frac{q_n(t_0)}{c}\right)\right). \quad (9)$$

It can be seen that the received signal differs from the emitted signal by a time delay $\boldsymbol{q}_n(t_0)/c$ and a time stretch $c/(c - v\cos\alpha_n)$.

As it is assumed that the position of the source is known at the time t_0, the time delay between the received and the emitted signal can be calculated,

$$\Delta T_n(t_0) = \frac{q_n(t_0)}{c}. \quad (10)$$

Let $x_n(t')$ be the portion of every receiver signal $y_n(t_0+t')$, originating when the source was at position $p(t_0 + t)$, and lasting for Δt:

$$x_n(t') = y_n\big(t_0 + t' + \Delta T_n(t_0)\big), \quad (11)$$

where $0 \le t' \le \Delta t$.

Combining equations (11) and (9),

$$\begin{aligned} x_n(t') &= s_e\left(t_0 + \frac{c}{c - v\cos\alpha_n}\left(t' + \Delta T_n(t_0) - \frac{q_n(t_0)}{c}\right)\right) \\ &= s_e\left(t_0 + \frac{c}{c - v\cos\alpha_n}t'\right). \end{aligned} \quad (12)$$

The signal portion $x_n(t')$ where $0 \le t' \le \Delta t$ doesn't correspond exactly to the source signal emitted from t_0 to $t_0 + t$. The difference can be studied by evaluating the received signal value at the beginning and end of the interval $t' \in [0, \Delta t]$,

$$x_n(0) = s_e(t_0), \quad (13)$$

$$x_n(\Delta t) = s_e\left(t_0 + \frac{c}{c - v\cos\alpha_n}\Delta t\right). \quad (14)$$

It can be seen from equation (14) that $x_n(\Delta t) \neq s_e(t_0+\Delta t)$. There is an extra (or missing if $\cos\alpha_n < 0$) portion of the signal in $x_n(t')$ for $0 \le t' \le \Delta t$ that corresponds to the portion of signal emitted by the source from $t_0 + \Delta t$ to $t_0+\Delta t\, c/(c-v\cos\alpha_n)$. The length of this signal portion differs depending on the receiver since it is a function of α_n. However, if $v = ct$, and the displacement of the source is considerably smaller than the distance between the source and the receiver during $\Delta t\, c/(c - v\cos\alpha_n)\forall\alpha_n$, then the Doppler stretch is constant. Therefore, the frequency content of all the signals is not modified by this signal portion, provided that the source signal is stationary.

In summary, the propagation delay from the source to every receiver at time t_0 can be calculated using equation (10). And the signal portion that was emitted within t_0 and $t_0 + \Delta t$ can be expected to correspond to the received signal $y_n(t_0 - t' + \Delta T_n(t_0))$ for $0 \le t' \le \Delta t$ and both signals just differ in the time stretching due to the Doppler Effect. Therefore, the received signals differ from each other in the same way. The relative Doppler dilation can be calculated (section 3), and there is a geometric relation between the relative Doppler dilation and the position and velocity of the source, thus these two parameters can be calculated if enough receivers are available (section 4). The new position of the source is assigned to the time instant $t_0+\Delta t/2$ and this allows for repeating the process by calculating the time in which the sound signal is received at each receiver for that new source position. Figure 2 shows a block diagram of the whole process. This procedure is valid as long as the following assumptions are valid:

1. The position $p(t_0)$ of the source is known at a certain instant t_0.
2. The moving source speed is a constant vector during the time interval $\Delta t\, c/(c - v\cos\alpha_n)\ \forall\alpha_n$.
3. The displacement of the source during $\Delta t\, c/(c - v \cdot\cos\alpha_n)$ is considerably smaller than the distance between the source and the receiver: $v\,\Delta t\, c/(c - v\cos\alpha_n) \ll q_n(t_0)\ \forall\alpha_n$.

3. Calculation of the relative Doppler dilation between two receivers

From equation (12) it can be seen that the signal portion $x_n(t')$ is time stretched with respect to the emitted signal $s_e(t')$. The time stretching due to the Doppler Effect in equation (12) can be expressed in the frequency domain as

$$X_n\left[f\left(\frac{c}{c - v\cos\alpha_n}\right)\right] = S_e(f). \quad (15)$$

The Doppler Effect will cause the frequency spectra obtained from the microphones to appear expanded or compressed in comparison with the original. This expansion or compression can be understood as if the frequency axis was rescaled by a factor δf_n. The rescaled frequency is going to be named f_n so that

$$X_n(f_n) = X_n(f \cdot \delta f_n) = S_e(f). \quad (16)$$

The term δf_n is going to be called the Doppler dilation and can be expressed as

$$\delta f_n = \frac{c}{c - v\cos\alpha_n} \quad (17)$$

or

$$\delta f_n = \frac{c}{c - \boldsymbol{v}\cdot\boldsymbol{u}_n}, \quad (18)$$

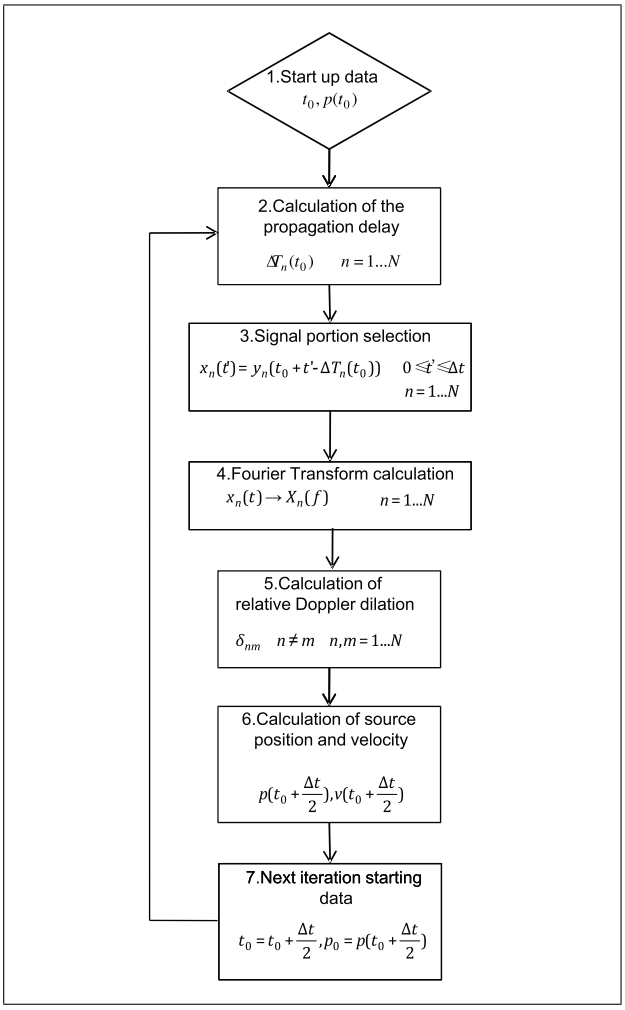

Figure 2. Block diagram showing the different stages in the signal processing algorithm.

where $\boldsymbol{v}$ is the velocity of the moving source, $\boldsymbol{u}_n$ is an unitary vector pointing from the source to the receiver n.

It is not possible to infer the absolute Doppler frequency dilation in each receiver (δf_n) without previous knowledge of the source velocity, or the shape of the original spectrum. However, as the received signal in two different microphones can be measured and compared, the relative Doppler dilation between them can be obtained instead of the absolute Doppler dilation in each receiver. The definition of the relative Doppler shift between two receivers is

$$\delta f_{nm} = \frac{f_n}{f_m} = \frac{\delta f_n}{\delta f_m}. \tag{19}$$

In SONAR and RADAR disciplines, matched filters are used to compare the transmitted signal ($x_t(t)$) and the received ($x_r(t)$). The received signal is basically the transmitted signal delayed by a certain amount of time τ_{tr} and affected by the Doppler dilation δf_{tr} in the case of moving sources. The output of the matched filter, in the case of approaching the Doppler Effect as a Doppler stretch instead of a Doppler shift, is given by the Wide Band Cross Ambiguity Function [14, 15]

$$\chi(\tau, \delta f) = \frac{1}{\sqrt{\delta f}} \int_{-\infty}^{+\infty} x_t(t)\, x_r\left(\frac{t}{\delta f_n} - \tau\right) \mathrm{d}t. \tag{20}$$

This output is a two dimensional function of the parameters τ and δf. The Wide Band Cross Ambiguity Function can also be obtained by performing the inverse Fourier Transform of the Cross-Spectra of the two signals,

$$\chi(\tau, \delta f) = \frac{1}{\sqrt{\delta f}} \int_{-\infty}^{+\infty} X_t(f)\, X_r^*(f\, \delta f_n)\, \mathrm{e}^{2\pi\, \tau\, \delta f}\, \mathrm{d}f. \tag{21}$$

The Wide Band Cross Ambiguity function has a global maximum at $\delta f = \delta f_{tr}$ and $\tau = \tau_{tr}$. Unlike in SONAR or RADAR systems, in the present situation the time delay between the signals of two receivers is compensated in steps 2 and 3 of the method (Figure 2). Then, $\tau_{nm} = 0$ and the signals can be compared using a one dimensional ($\tau = 0$) version of equation (21),

$$\chi(\tau, \delta f) = \frac{1}{\sqrt{\delta f}} \int_{-\infty}^{+\infty} X_n(f)\, X_m^*(f\, \delta f_n)\, \mathrm{d}f. \tag{22}$$

Equation (22) will have a global maximum at $\delta f = \delta f_{nm}$. Moreover, if there is no time delay between the signals $X_n(f)\, X_m^*(f\, \delta f_n) = |X_n(f)\, X_m^*(f\, \delta f_n)|$, then

$$\chi(\tau, \delta f) = \frac{1}{\sqrt{\delta f}} \int_{-\infty}^{+\infty} \left| X_n(f)\, X_m^*(f\, \delta f_n) \right| \mathrm{d}f. \tag{23}$$

The use of the one dimensional version of the Ambiguity Function (equation 23) instead of the 2D version (equation 21) reduces the computational complexity of the algorithm. Furthermore, given the properties of the Cross Spectra shown in equation (23), the influence of uncorrelated background noise is negligible.

Note that, in practice the discrete version of equation (23) will be used. As only $X_m(f)$ is measured, a resampling of the frequency axis is needed. In the results showed later in the text, the resampling is made by linear interpolation.

4. Geometric relations

Let $\boldsymbol{p}$ be the position vector of the moving source in the spatial reference $\{o, \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$, and r_n be the position vector of the receiver n.

Given this geometry it can be written

$$\boldsymbol{v} \cdot \boldsymbol{u}_n = \boldsymbol{v} \cdot \frac{\boldsymbol{q}_n}{q_n}. \tag{24}$$

The combination of equation (18) and equation (24) gives

$$\delta f_n = \frac{c}{c - \boldsymbol{v} \cdot \boldsymbol{q}_n / q_n}. \tag{25}$$

Therefore, the relative Doppler shift between two receivers can be written as

$$\delta f_{nm} = \frac{c - \boldsymbol{v} \cdot \boldsymbol{q}_m / q_m}{c - \boldsymbol{v} \cdot \boldsymbol{q}_n / q_n}. \tag{26}$$

Considering the whole set of receivers, and taking one of them as a reference, a non linear system of equations can be obtained by writing equation (26) for the reference microphone paired with any other. The resulting system will have the three components of $\boldsymbol{v}$ and the three components of $\boldsymbol{p}$ as unknown variables, thus, six or more equations are needed in order to solve the system of equations. Seven is therefore the minimum number of receivers required for this locating method to work.

5. Results

The experimental setup designed in order to test the performance of the method consists of a mobile device running at constant speed (around 3.5 m/s) along a linear track inside a semi-anechoic chamber, with seven microphones aligned parallel to the track. The distance between two adjacent microphones is 0.56 m. Also, a set of 17 infrared light barriers are placed along the track, when the mobile device passes between them, an electric pulse is registered and the position of the source is confirmed at that time instant. Also the speed can be obtained from position differences. The distance between two adjacent infrared light barriers is 0.21 m. Figure 4 shows the whole experimental setup.

A small loudspeaker (5 cm diameter) is mounted on the mobile device so the acoustic signal emitted can be controlled. Tests have been performed when the loudspeaker emits a 10 kHz single tone, and also white noise. The seven microphones recorded the acoustic pressure signal of the moving source running along the track at a sampling rate of 50 kHz, which is enough to avoid aliasing at the emitted frequency range. Only the recorded signal corresponding to the effective track segment was considered for further processing, this signal segment is the one contained within the pulse of the first and the last light barriers.

The goal of this test is to measure the position and velocity of the mobile sound source from the acoustic signal received at each microphone. The result can be compared with the real position and speed of the moving source given by the light barriers. The source is released about a meter before the beginning of the effective track, i.e. the position of the first light barrier $\boldsymbol{p}(t_0)$. The light barrier records the time t_0 at which the source is at $\boldsymbol{p}(t_0)$ so that assumption 1 in section 2 is upheld. Also assumption 2 is upheld since the acceleration period was finished when the mobile source reached the effective segment of the track so its speed is approximately constant. The procedure depicted in Figure 2 is used to locate the source taking $\Delta t = 0.082\,\mathrm{s}$ even though assumption 3 is not upheld. It can't be stated that $v\,\Delta t\,c/(c - v\cos\alpha_n) \ll \min(q_n(t_0))$ since $v\,\Delta t = 3.5\,\mathrm{m/s} \cdot 0.082\,\mathrm{s} = 0.29\,\mathrm{m}$ and $\min(q_n(t_0)) = 0.82\,\mathrm{m}$ (which is the perpendicular distance between the microphones and the track). Therefore, the method is tested in unfavorable conditions.

Choosing a smaller Δt or a larger distance from the microphones to the track in order to fulfil assumption 3 will lead to other problems in the current set up: On one hand, the Doppler dilation is given by equation (17) and, as in this case $v \ll c$, there will be small differences of Doppler dilation between the microphones due to the angle α_n. The system may not be able to detect the Doppler dilation differences if the range of α_n is small, therefore, microphones need to be placed close to the track. The thick lines in Figure 5 show the relative Doppler dilation taking microphone 1 as the reference (it takes the mobile device about 1 s to run along the effective track). On the other hand, the length of the signal portion is Δt, so the frequency spectra resolution will be $\Delta f = 1/\Delta t$. For a single tone 10 kHz signal, the estimation of the relative Doppler dilation will be as shown in equation (19),

$$\hat{\delta} f_{nm} = \frac{\hat{f}_n}{\hat{f}_m} = \frac{f_n \pm \Delta f/2}{f_m \pm \Delta f/2}. \tag{27}$$

Therefore, the error made in the estimation of δf_{nm} is

$$E\,\delta f_{nm} = \delta f_{nm} - \hat{\delta} f_{nm}. \tag{28}$$

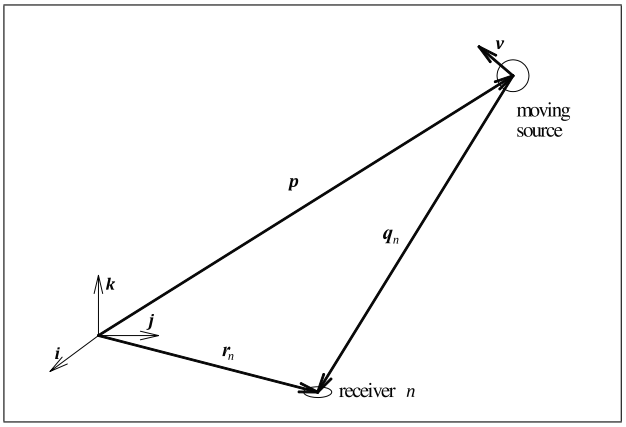

Figure 3. Diagram showing the vectors involved in the definition of the geometry of the problem. $\boldsymbol{p}$ is the position vector of the moving source in the spatial reference $\{o, \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$, and $\boldsymbol{r}_n$ is the position vector of the receiver n.

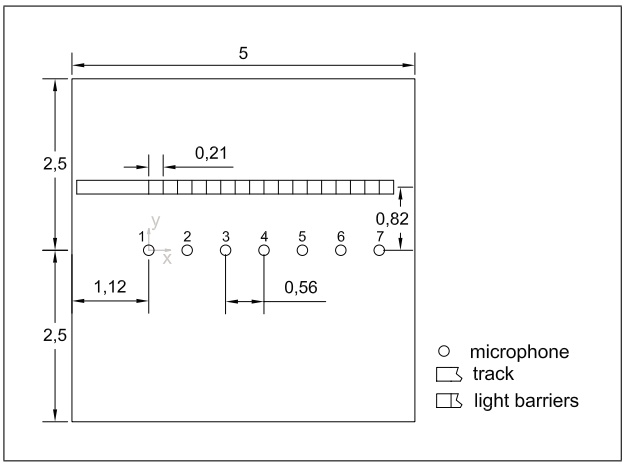

Figure 4. Sketch of the set up used both in the simulation and in the experimental test.

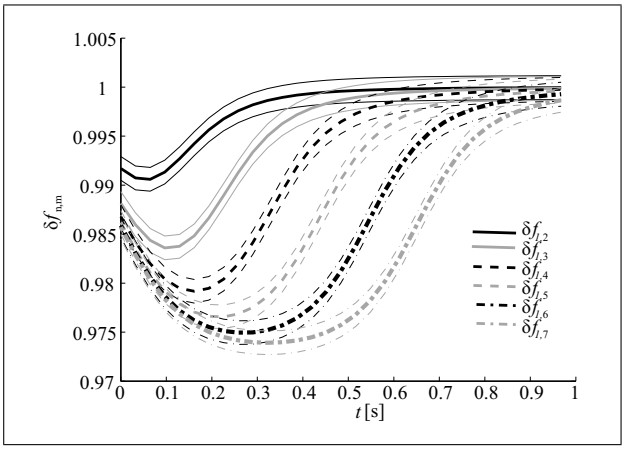

Figure 5. The relative Doppler dilation taking microphone 1 as the reference (thick lines), and margin of error in the estimation of δf_{nm} (thin lines).

Combining equations (19) and (27) into equation (28),

$$E\,\delta f_{nm} = \frac{f_n}{f_m} - \frac{f_n \pm \Delta f/2}{f_m \pm \Delta f/2} = \pm\frac{\delta f_{nm} + 1}{\Delta t\, f_{nm} \pm 1/2}. \qquad (29)$$

The thin lines in Figure 5 depict the margin of error in the estimation of δf_{nm} for the present case when $\Delta t = 0.082$ s.

Eventually the difference between the relative Doppler dilations is seen to be even smaller than the estimation errors. This could be solved by choosing a larger δt to reduce the error as it is seen in equation (29). However, this could cause assumption 3 not to hold true at all.

5.1. Simulation

The system is first tested by simulation in order to have a reference of the performance of the method without assumption 3 upheld, and also without measurement errors. The speed of the sound source moving along the linear track is 3.5 m/s and it emits a 10 kHz sinusoid signal. The signal recorded by the 7 receivers while the source is moving is simulated according to equation (6) with $s_e(t_0 + t) = \sin(2\pi\, 10000(t_0 + t))$.

In general, two main error sources can affect the tracking method,

- Inaccurate knowledge of the previous source position can lead to errors in estimating the propagation delay (step 2 in Figure 2).
- Numerical error in solving a non-linear system of equations (equation26) introduces inaccuracies in the estimated source position (step 6 in Figure 2).

As the described method is iterative, these sources of error can cause the total error to grow in time.

To isolate and evaluate the error introduced only by solving the non-linear system of equations (least squares in this example), the procedure in Figure 2 is applied, but at step 2, the real position of the source is always considered instead of the estimated one. In this way, source tracking errors do not propagate and only the error from solving the non-linear system of equations appears. The simulation results from this numerical error only case are

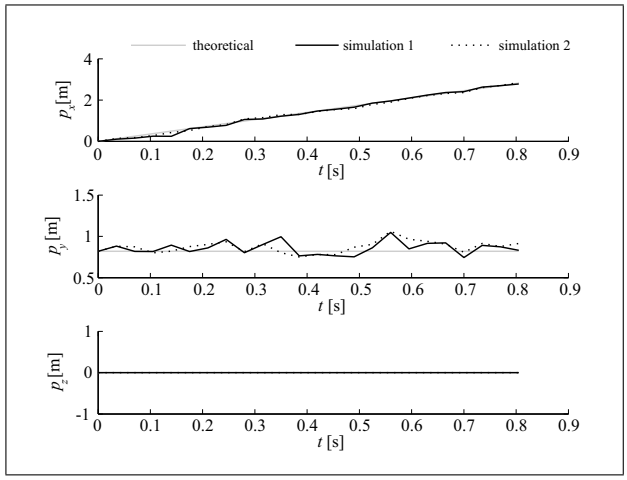

Figure 6. A comparison of position results using three methods. "Simulation 1" contains numerical errors associated with the iterative solution of the non-linear set of equations using a least squares method. "Simulation 2" is the original simulation and contains both the numerical errors in "Simulation 1" as well as errors in the source position. "Theoretical" refers to the real position of the source.

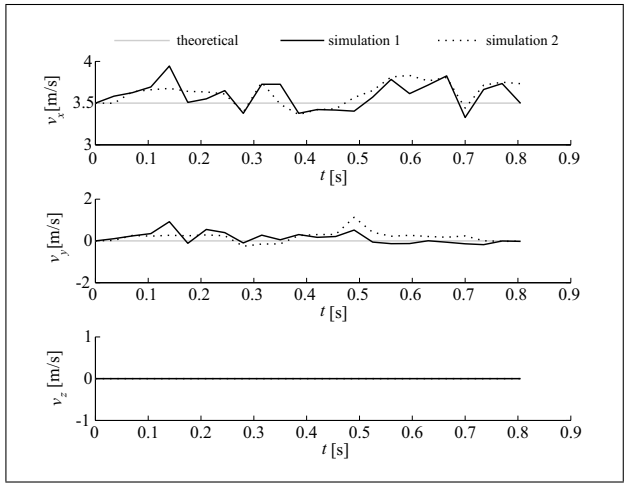

Figure 7. A comparison of velocity results using three methods. "Simulation 1" contains numerical errors associated with the iterative solution of the non-linear set of equations using a least squares method. "Simulation 2" is the original simulation and contains both the numerical errors in "Simulation 1" as well as errors in the source position. "Theoretical" refers to the real position of the source.

given as Simulation 1 in Figures 6 and 7. The simulation results from the original procedure are also shown in Figures 6 and 7 as Simulation 2.

In Step 6 of Figure 2, the system of equations introduced in equation (26) is solved using an iterative method of least squares, thus it is necessary to provide a starting position and velocity vector. In the case of Simulation 1 the starting position and speed are the theoretical ones, in the case of Simulation 2 they are estimated from the previous position assuming that the previous estimated velocity remains constant.

Table I. Maximum position and velocity errors found in simulation 1 and simulation 2.

	simulation 1	simulation 2
p error [m]	0.25	0.26
v error [m/s]	1.13	1.02

In Figure 6 and Figure 7, the simulated results show a clear tendency to follow the theoretical results. Comparing the results of Simulation 1 and Simulation 2 it is seen that having a slightly biased estimation of the previous position doesn't increase the error in general. Also, in the case of Simulation 2, the error doesn't tend to increase with time in spite of the tracking method being iterative. Therefore it seems that the fact of solving the equation system using a least square approximation minimizes the error introduced by a slightly biased estimation of the position. This can be explained because a biased estimation of the source position would lead to a biased signal portion selection (steps 2 and 3 in Figure 2). An inaccurate signal portion selection would lead to a bad estimate of δf_{nm}. However, if the selection is just slightly biased, assumption 2 and 3 would uphold and δf_{nm} will not be affected. Assuming then that the error in set of δf_{nm} is negligible, the least square method approximates the theoretical position even if the previous estimation was slightly biased. Table I shows the maximum error in position and velocity in simulations 1 and 2 defined as

$$p^{\text{error}} = \Big[(p_x^{\text{simulated}} - p_x^{\text{theoretical}})^2 + (p_y^{\text{simulated}} - p_y^{\text{theoretical}})^2 + (p_z^{\text{simulated}} - p_z^{\text{theoretical}})^2\Big]^{1/2}, \tag{30}$$

$$v^{\text{error}} = \Big[(v_x^{\text{simulated}} - v_x^{\text{theoretical}})^2 + (v_y^{\text{simulated}} - v_y^{\text{theoretical}})^2 + (v_z^{\text{simulated}} - v_z^{\text{theoretical}})^2\Big]^{1/2}. \tag{31}$$

If $\Delta t = 0.082$ s and the source speed is 3.5 m/s, during this time the source will have travelled $\Delta d = 0.29$ m, therefore there is an inevitable uncertainty of $\pm\Delta d$ on the position of the source. It can be seen that the p error is of the same order of magnitude. The biggest error appears when the simulated value of v is 30% of the source theoretical speed.

It is also performed a set of simulations starting with a biased initial position of the source in order to test the robustness of the method. Errors of +0.01 m, +0.02 m, +0.05 m, +0.1 m, +0.2 m are introduced in the x-coor di nate of the initial position. Figure 8 and Figure 9 show the obtained result.

Figure 8 and Figure 9 show that, in general, the error introduced in the initial position doesn't increase with time. Table II lists the maximum errors found during the source displacement in every simulation (using equations 30 and 31), together with the mean error values. It comes out that the maximum error and the mean error increase

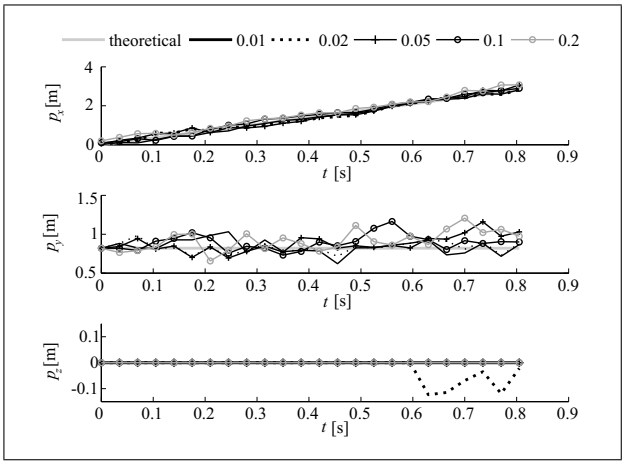

Figure 8. Position results obtained introducing errors of +0.01 m, +0.02 m, +0.05 m, +0.1 m, +0.2 m in the x-coordinate of the initial position.

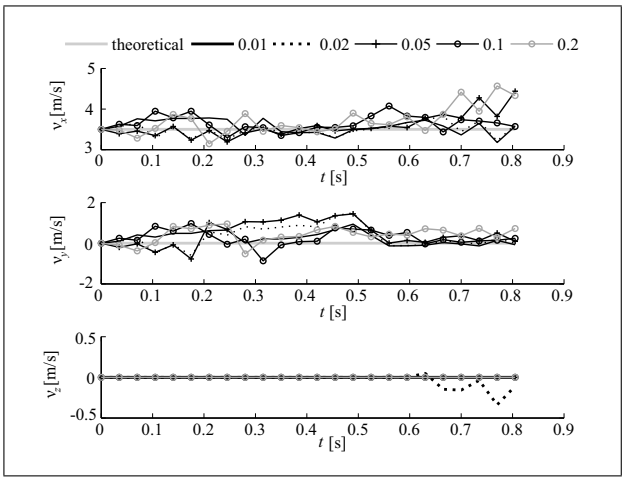

Figure 9. Position results obtained introducing errors of +0.01 m, +0.02 m, +0.05 m, +0.1 m, +0.2 m in the x-coordinate of the initial position.

with the error in the initial position. For initial errors smaller than +0.005 m, the maximum error in position is always smaller than $\Delta d = 0.29$ m, which is the inevitable uncertainty, and the mean error is similar. It seems that the method is robust for initial position errors smaller than the diameter of the source (diameter of the speaker 0.05 m). For initial errors up to the same order of magnitude as Δd, the mean error remains smaller than Δd, although important maximum errors are detected. However, note than an initial error of +0.2 m equals to the 25% of the distance between the track and the receivers (0.82 m) in that case, therefore it is a big initial error. As smaller is the ratio between the initial position error and the distance between the source and the receiver, smaller is the influence of the initial position error. This is because the difference in propagation time from the source to the receiver will be negligible if the initial position error is negligible compared to the distance between source and receiver.

Table II. Maximum and mean position and velocity errors found when there is no error in the initial position (+0), and when there is an error of +0.01 m, +0.02 m, +0.05 m, +0.1 m, +0.2 m in the x-coordinate of the initial position.

		+0	+0.001	+0.002	+0.005	+0.1	+0.2
mean	p error [m]	0.095	0.10	0.12	0.15	0.15	0.23
	v error [m/s]	0.31	0.35	0.51	0.68	0.42	0.61
max	p error [m]	0.25	0.26	0.23	0.41	0.36	0.52
	v error [m/s]	1.13	0.94	1.50	1.44	1.06	1.10

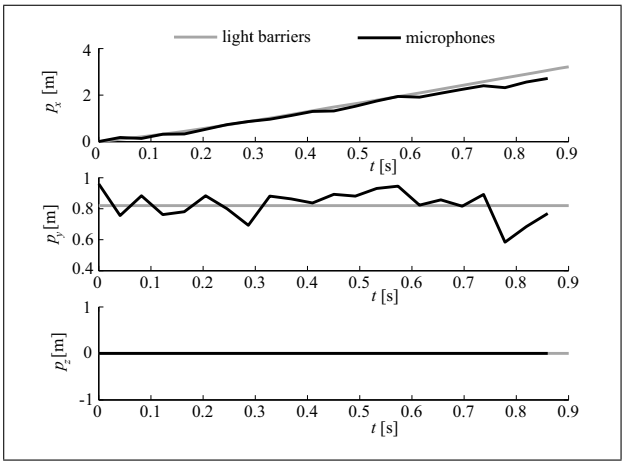

Figure 10. Comparison between the time history for the three components of the position obtained from the set of light barriers and obtained by the microphone system. The source is emitting a single tone of 10 kHz.

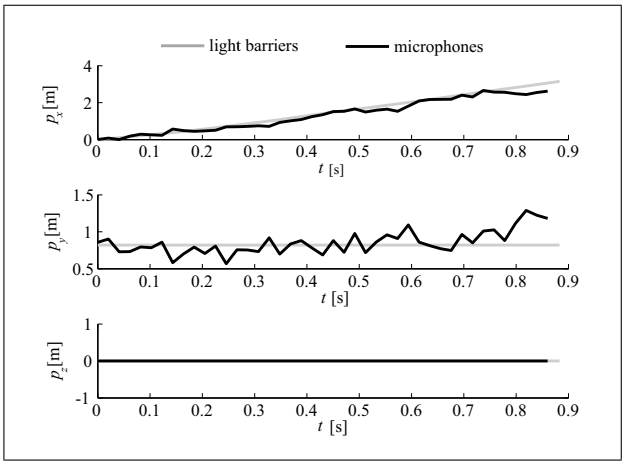

Figure 12. Comparison between the time history for the three components of the position obtained from the set of light barriers, and obtained by the microphone system. The source is emitting white noise.

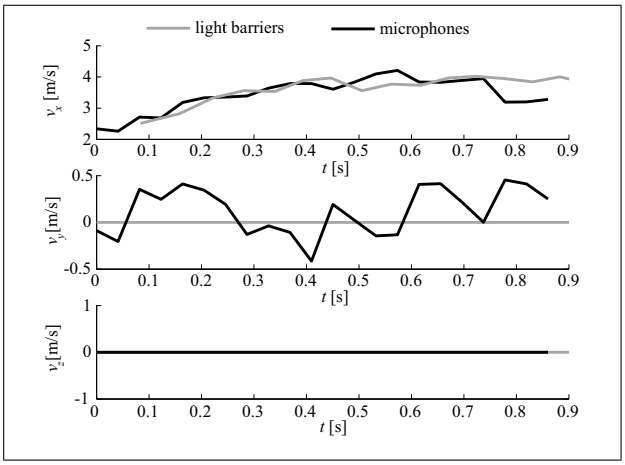

Figure 11. Comparison between the time history for the three components of the velocity obtained from the set of light barriers and obtained by the microphone system. The source is emitting a single tone of 10 kHz.

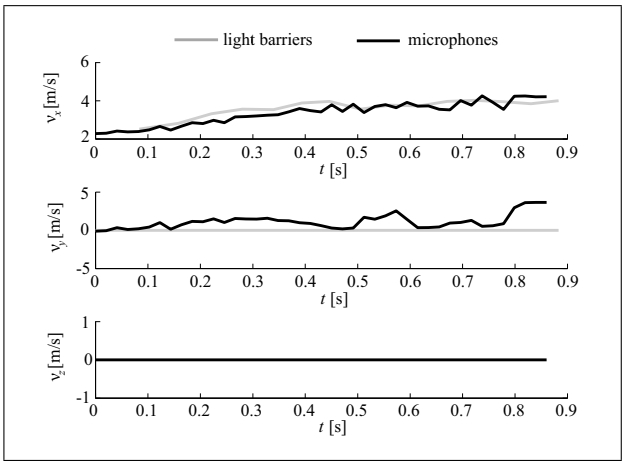

Figure 13. Comparison between the time history for the three components of the velocity obtained from the set of light barriers and obtained by the microphone system. The source is emitting white noise.

5.2. Experimental test

In the first experimental test, the loudspeaker emits a 10 kHz sinusoid, in the second test it emits white noise. Figure 10 and Figure 11 compare the position values p_x, p_y, and p_z the speed values v_x, v_y, v_z, given by the light barriers system and by the microphone measurement system developed here, in the case of a 10 kHz sinusoid. Using the light barriers system, the time that it takes the source to travel from one light barriers to another can be measured and the source's speed calculated.

Figure 12 and Figure 13 compare the position values p_x, p_y, and p_z the speed values v_x, v_y, v_z given by the light barriers system and by the microphone measurement system in the case of the source emitting white noise.

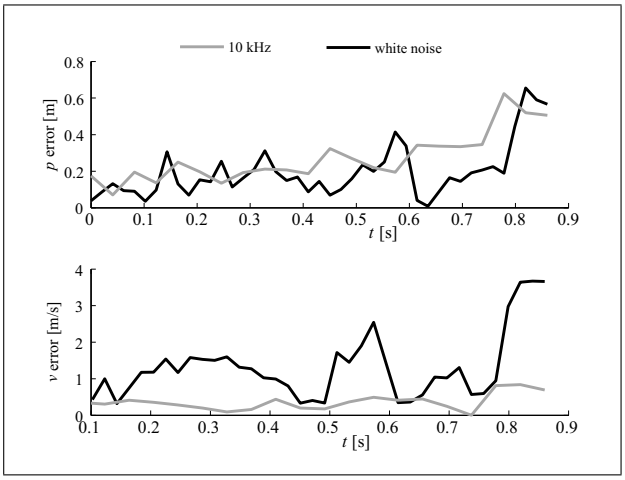

Figure 14. Error found in the position estimation as a function of time (top), error found in the estimation of the speed as a function of time (bottom).

It follows from Figure 10 to Figure 13 that the microphone system follows the source movement. At that point the errors are not only due to the violation of assumption 3, but also due to measurement errors. Figure 14, shows the evolution of the error as a function of time computed as said in equation (30) and equation (31) taking the position and speed provided by the light barriers as the theoretical.

Table III lists the maximum and mean error values. Regarding the test with 10 kHz, comparing Table III with the first column of Table II, the velocity errors (max and mean) are of the same order of magnitude. The mean position value is higher in the experimental test but still smaller than Δd, the main difference is the maximum position error. The maximum error in position happens at the end of the recording (see Figure 14), just at the moment in which it is more difficult to obtain a good estimation of δf_{1m} (see Figure 5) because of the unfavorable set up. Excluding the data after 0.75 s the maximum error in position in the experimental test would be 0.34 m, which is similar to Δd.

In the case of white noise, the position errors are similar to those in the case of 10 kHz. The maximum position error again happens at the end of the recording. Excluding the data after 0.75 s the mean position error would be 0.16 m, again smaller Δd. The velocity errors are much larger in the case of white noise, this shows that the fact of not holding assumption 3 may have a greater effect in broadband noise. If assumption 3 is not held then $\dot{\alpha}_n \neq 0$, and thus the Doppler stretch is not constant during the considered portion of the signal. Therefore, the profile of the frequency spectra is not sharp but blurry affecting the performance of the 1D Ambiguity Function. Although, the errors in velocity are much larger than the position ones, it is not critical since the velocity can be derived from the position.

As it has been said before, an example of application of the present method is aircraft tracking by using a mesh of receivers spread around the airport. In that case, background noise in the measurement sites is expected, and is also expected that this background noise is uncorrelated between the microphones. To evaluate the robustness of the method in this situation, the results obtained when uncorrelated background noise is added to the microphones signals are shown in Table IV. The background noise used in this test is Gaussian white noise. It can be seen that the mean position error is still close to Δd up to a signal-to-noise ratio of +3 dB. No relevant differences between the cases of +6 dB and +3 dB are found. It is also important to note that no relevant differences are found when the source emits a 10 kHz sine signal or white noise, illustrating that the method is independent of the signal characteristics.

Table III. Maximum position and velocity errors found in the experimental tests with a single tone of 10 kHz and white noise.

		10 kHz	White noise
mean	p error [m]	0.27	0.20
	v error [m/s]	0.36	1.32
max	p error [m]	0.62	0.65
	v error [m/s]	0.84	3.67

Table IV. Maximum position and velocity errors found in the experimental tests with a single tone of 10 kHz and white noise adding background noise.

		10 kHz	White noise
		SNR +6 dB	
mean	p error [m]	0.30	0.27
	v error [m/s]	1.40	1.10
max	p error [m]	0.78	0.60
	v error [m/s]	3.06	4.07
		SNR +3 dB	
mean	p error [m]	0.32	0.27
	v error [m/s]	1.18	1.15
max	p error [m]	0.81	0.61
	v error [m/s]	3.54	3.95

6. Conclusions

The acoustic moving sound source tracking system uses a minimum of 7 microphones, distributed in the acoustic influence area of the source. It is based on the relative Doppler dilation between a reference microphone and the remaining microphones, which depends on the position and velocity of the source. The method uses a 1D version of the Ambiguity Function to obtain the relative Doppler dilation between a pair of microphones. The method developed in this paper has the capability to track the source and measure its velocity with the following assumptions:

1. The position of the source is known at least at one random time instant.
2. The moving source speed is considered constant for short time intervals $\Delta t\, c/(c - v\cos\alpha_n)\ \forall\alpha_n$.
3. The source displacement during $\Delta t\, c/(c - v\cos\alpha_n)\ \forall\alpha_n$ is always negligible compared to the distance between the source and the receivers.

Simulations have shown the method to be robust against errors of the same order of magnitude than the dimension of the source in the initial, known source position, described above in Assumption 1. Furthermore, although the method relies on an iterative numerical scheme, the total error does not increase with time because the least square approximation minimizes the error introduced by a slightly biased estimation of the previous position.

Experimental comparisons are made using both single-frequency and white noise signals and show an increased error in comparison to the simulations results, especially at the end of the recording due to the unfavorable set up (violation of assumption 3). However, the errors remain comparable to the distance travelled by the source during the considered interval Δt, even at signal-to-noise ratios of up to +3 dB, showing the robustness of the method in the presence of background noise. Also, the results are similar in the case of broadband noise and single tone, illustrating that the performance of the method doesn't depend on the signal characteristics.

In summary, the tests show that if the three assumptions are valid, the method is capable to locate the source with a resolution comparable to the distance travelled by the source during the considered interval Δt.

References

[1] P. Svaizer, M. Matassoni, M. Omologo: Acoustic source location in a three-dimensional space using crosspower spectrum phase. CASSP, IEEE International Conference on Acoustics, Speech and Signal Processing, 1996, 921–924.

[2] J. E. Barger: Sonar systems. – In: Encyclopedia of acoustics, Chapter 49, vol.1. M. J. Crocker (ed.). John Wiley and Sons, Inc., 1997, ISBN 0–471–80465–7.

[3] G. C. Carter: Time delay estimation for passive sonar signal processing. IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-29. No. 3, June 1981.

[4] D. H. Fohnson, D. E. Dudgeon: Array signal processing. Prentice-Hall, ISBN 0-13-048513-6.

[5] K. Brensing, K. Linke, D. Todt: Sound source location by difference of phase, on a hydrophone array with small dimensions. J. Acoust. Soc. Am. **109** (2001) 430–433.

[6] H. Chen, J. Zhao: On locating low altitude moving targets using a planar acoustic sensor array. Applied Acoustics **64** (2003) 1087–1101.

[7] B. G. Fergusson, L. G. Criswick, K. W. Lo: Locating far field impulsive sound sources in air by triangulation. J. Acoust. Soc. Am. **111** (2002) 104–116.

[8] R. Blumrich, J. Altman: Medium-range localisation of aircraft via triangulation. Applied Acoustics **61** (2000) 65–82.

[9] B. G. Fergusson: Time-delay estimation techniques applied to the acoustic detection of jet aircraft transits. J. Acoust. Soc. Am. **106** (1999).

[10] S. Nagata, K. Furihata, T. Wada: A three-dimensional sound intensity measurement system for sound source identification and sound power determination by ln models. J. Acoust. Soc. Am. **118** (2005) 3691–3705.

[11] P. Rizzo, G. Bordoni, A. Marzani, J. Vipperman: Localization of sound sources by means of unidirectional microphones. Meas. Sci. Technol. **20** (2009) 055202 (12pp).

[12] T. S. Brandes, R. H. Benson: Sound source imaging of low-flying airborne targets with an acoustic camera array. Applied Acoustics **68** (2007) 752–765.

[13] J. I. Statman, E. R. Rodemich: Parameter estimation based on Doppler frequency shifts. IEEE Transactions on aerospace and electronic systems. Vol. AES-23, No.1, January 1987.

[14] K. W. Lo, B. G. Ferguson: Flight path estimation using frequency measurements from a wide aperture acoustic array. IEEE Transactions on Aerospace and Electronic Systems **37** (2001) 685–1694.

[15] Q. Jin, K. M. Wong, Z.-Q. Luo: The estimation of time delay and Doppler stretch of wideband signals. IEEE transactions on Signal Processing **43** (1995) 904–916.

[16] X. X. Niu, P. C. Ching, Y. T. Chan: Wavelet based approach for joint time delay and doppler stretch measurements. IEEE Transactions on Aerospace and Electronic Systems **35** (1999) 1111–1119.