Dual microphone system for 2D panoramic sound

Ashihara, Kaoru (AIST)

April 20, 2021

Contents

1 Preface **2**

2 Direction of propagation **4**

2.1 Spatial property - front-to-rear - 4

2.2 Spatial property - left-to-right - 5

3 Speed of propagation **8**

4 Binaural reproduction **9**

5 Conclusion **10**

1 Preface

In principle, a 3-point measurement is necessary to obtain 2-dimensional spatial information. Standard recording equipment, however, usually has only the left and right tracks. To use 3 or more microphones at the same time, therefore, some sort of special hardware setup is required. As can be seen in Fig. 1, 2 or more AD converters have to be synchronized to make a 3-track recording available. If 2-dimensional panoramic sounds can be captured by a 2-track recording, the system can be much simpler. In fact, a single stereo AD converter is all you need.

To this end, a new recording system for 2-dimensional panoramic sound was proposed. Figure 2 shows the outlook of the proposed system that consists of a pair of cardioid microphones. Since it has only 2 microphones, the system can be used with a simple 2-track recorder. For example, handy recorders such as 'ZOOM H6' and 'TASCAM DR-22WL' can be used for this purpose. A stereo microphone system for iOS devices, 'ZOOM iQ6' is also available. They are shown in Fig. 3.

As mentioned earlier, the system consists of a pair of cardioid microphones. In this document, they are referred to as Microphone A and Microphone B, and their outputs are referred to as Signal A and Signal B, respectively. As shown in Fig. 4, Microphone A and Microphone B are aligned side by side on the X-axis. In this figure, the beam directions of Microphone A and Microphone B are 0° and 180°, respectively, on the horizontal plane.

In principle, spatial properties of a wave at a given point can be specified by giving the direction and speed of propagation for each frequency bin. How they can be estimated when only Signal A and Signal B are available will be described in the following sections.

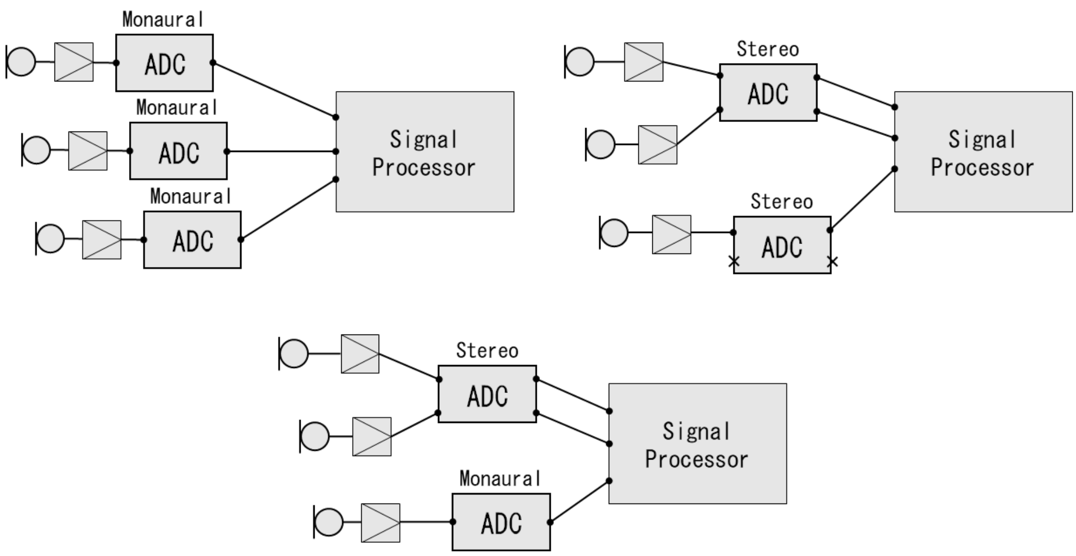

Fig. 1 Configurations of a 3-track recording
A 3-track recording usually requires either of 3 monaural AD converters (Top left), 2 stereo AD converters (Top right), or a stereo AD converter and a monaural AD converter (Bottom).

Fig. 2 Dual microphone system
The proposed dual microphone system consists of a pair of cardioid microphones.

Fig. 3 Handy recording devices
Handy recorders 'ZOOM H6' (Left) and 'TASCAM DR-22WL' (Center) and a stereo microphone system 'ZOOM iQ6' attached to a mobile device (Right) are shown. They can be used for capturing 2-dimensional panoramic sound.

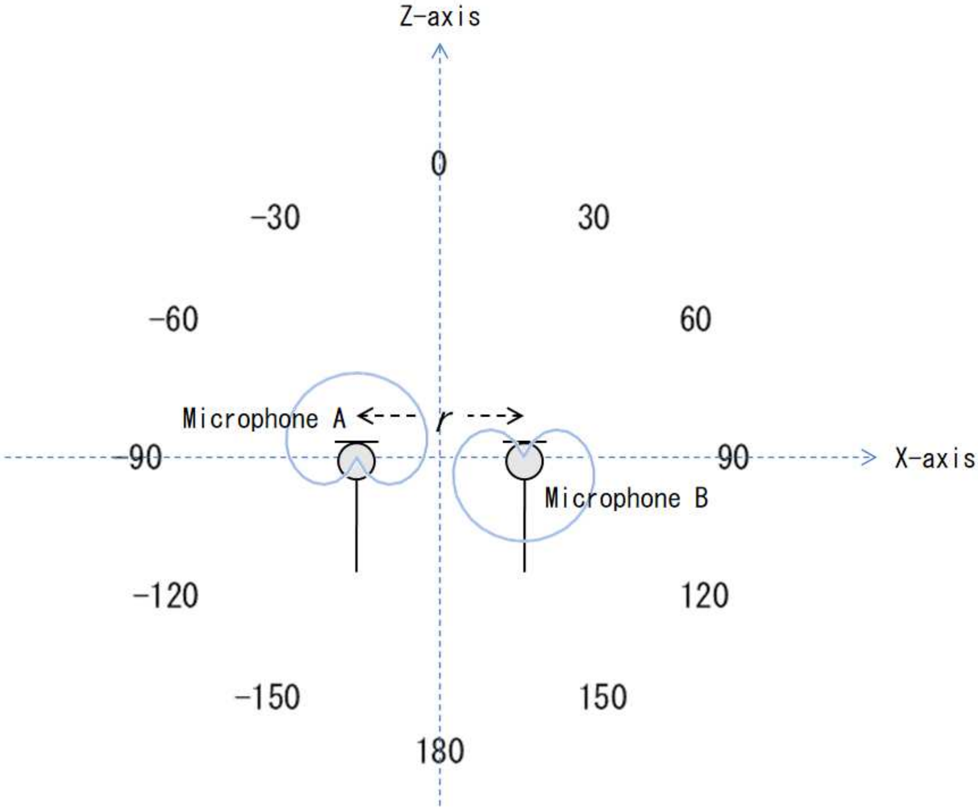

Fig. 4 Dual microphone system for 2D panoramic sound
A top view of the system is schematically drawn. The system consists of a pair of cardioid microphones arranged side by side on the X-axis. 'r' represents the distance between the microphones.

2 Direction of propagation

2.1 Spatial property - front-to-rear -

Since Microphone A is a cardioid microphone, its sensitivity is higher for the frontal direction (0° direction) and lower for the rear direction (180° direction) in the case of Fig. 4 and Microphone B faces the opposite direction. Therefore, the inter-channel level difference (level difference between Signal A and Signal B) contains a clue of one-dimensional (front-to-rear) spatial information.

If the beam pattern of Microphone B is an ideal cardioid, its sensitivity is NULL for the direction of 0°. The signal to noise ratio (S/N) of Signal B, therefore, is very poor around the direction of 0° in Fig. 4 and it causes a problem in calculation of the inter-channel phase difference that will be discussed later. When a sound source is close to the NULL direction, for instance, the direct sound from the source is suppressed and reflected sound from the opposite direction becomes dominant as can be seen in the right illustration of Fig. 5. If the microphone is omnidirectional instead of cardioid, the direct sound is most dominant as shown in the left illustration. As a result, there can be significant difference between phases measured by an omnidirectional microphone and a cardioid microphone of which the former is supposed to be the real value at the point. To reduce this difference, the beam direction of the microphones shown in Fig. 2 is substantially elevated from the horizontal plane (X-Z plane in Fig. 4) so that there is no NULL in the horizontal plane. This makes the sensitivity of the microphones not too low for any directions (in the horizontal plane). Since there still is front-to-rear sensitivity

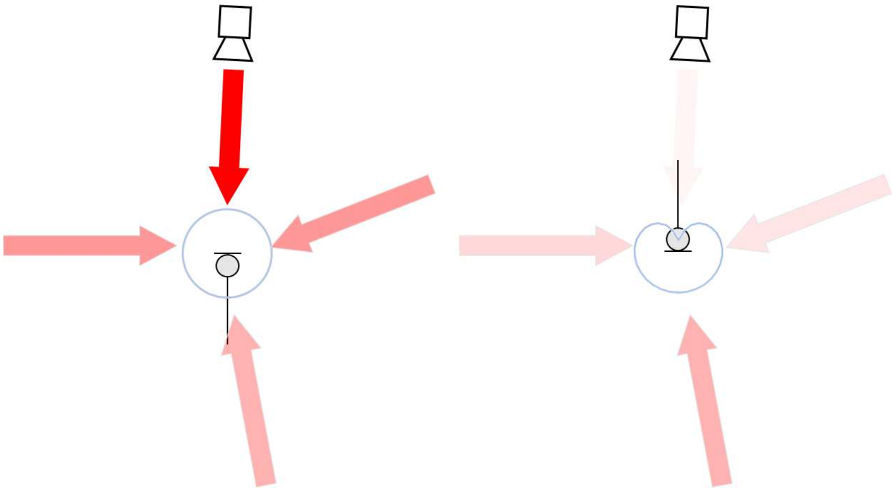

Fig. 5 Influence of reverberation
In the case of an omnidirectional microphone (Left), direct sound is most dominant. In the case of a cardioid microphone (Right), reflection from the rear is most dominant since direct sound is suppressed. In such cases, the phase of the components measured by the cardioid microphone may be less accurate.

difference, accuracy in the phase measurement can be poor especially in an environment with high reverberation.

Since the sensitivity of the cardioid microphones depends on direction, the inter-channel level difference as a function of direction will be like what is shown in the right graph of Fig. 6. It has its positive maximum (Signal A $>$ Signal B) at the 0° direction and the negative maximum (Signal A $<$ Signal B) at the 180° direction. The inter-channel level difference, therefore, can be used as a clue for the front-to-rear spatial information.

Figure 7 shows the inter-channel level differences of a handy recorder 'TASCAM DR-22WL' actually measured in a room with moderate reverberation. It can be seen that the level difference tended to be large in the 0° direction and 180° direction and small in the ±90° directions regardless of the frequency band of the sound indicating that the inter-channel level difference can be used as a clue for the front-to-rear spatial information.

Since the polar pattern of a cardioid microphone is symmetrical around its beam-axis, the inter-channel level difference has no clue of the left-to-right lateralization. For example, in the right graph of Fig. 6, when the inter-channel level difference is 0 dB, the sound can be coming either from the direction of 90° or that of −90°.

2.2 Spatial property - left-to-right -

Microphone A and Microphone B are aligned side by side on the X-axis and there is some distance between them. This will make the inter-channel phase difference that depends on the left-to-right lateralization of the sound. If the inter-channel phase difference has its positive maximum at the −90° direction and its negative maximum at the 90° direction or vice versa as shown in the right graph of Fig. 8, it can be a clue of one-dimensional (left-to-right) spatial information.

Figure 9 shows the inter-channel phase difference of TASCAM DR-22WL actually measured in a room with moderate reverberation. Although it changed irregularly in the frequency band below 1,500 Hz (top left), the phase difference was roughly similar to

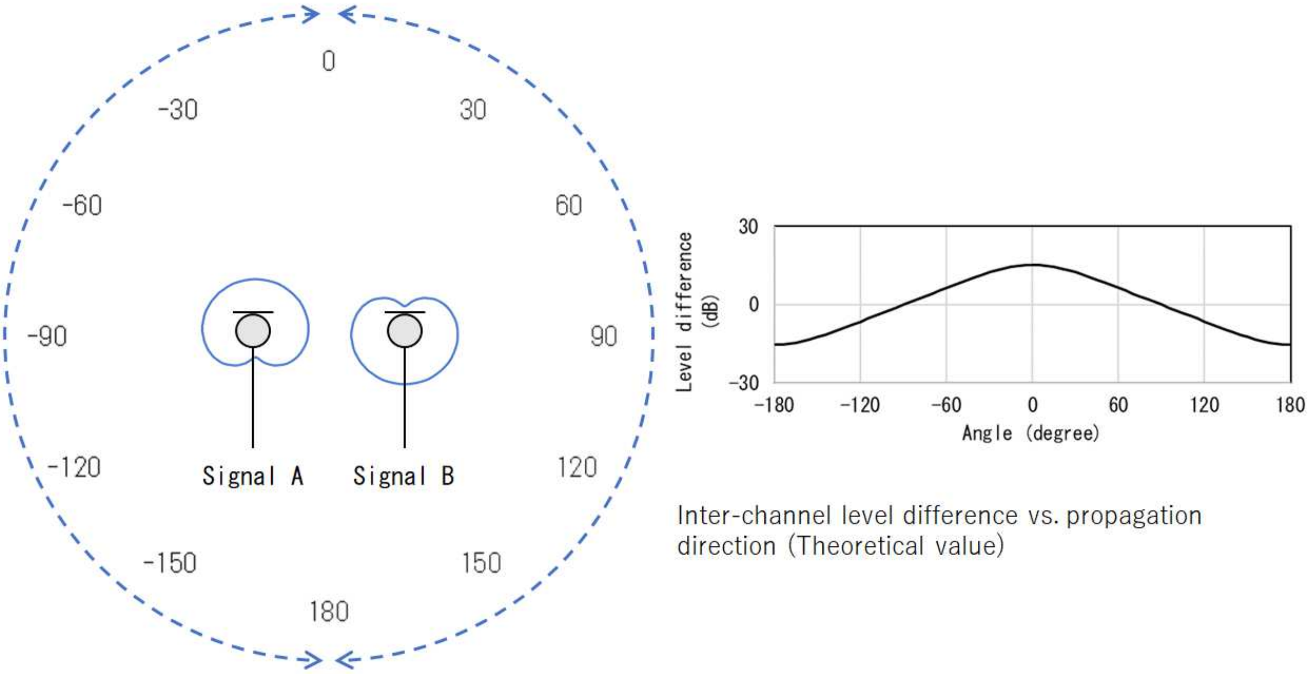

Fig. 6 Inter-channel level difference
A top view of the system is on the left. Inter-channel level difference (theoretical figure) is plotted on the right as a function of direction

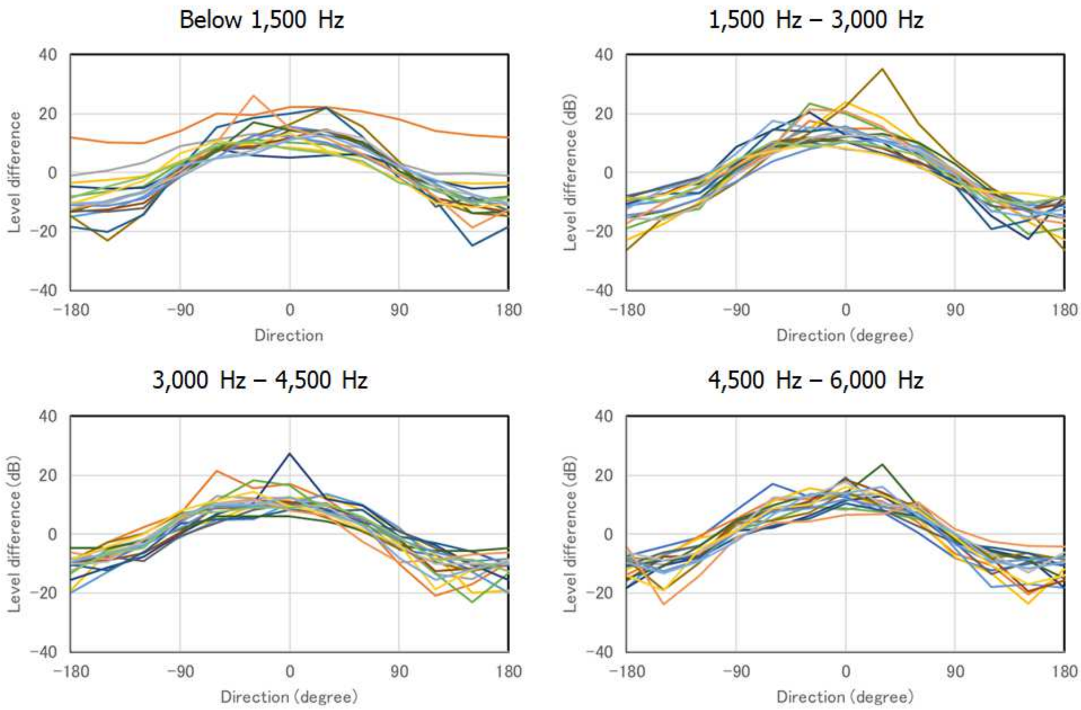

Fig. 7 Inter-channel level difference
Inter-channel level difference of TASCAM DR-22WL is plotted as a function of direction. Each line corresponds to each frequency bin.

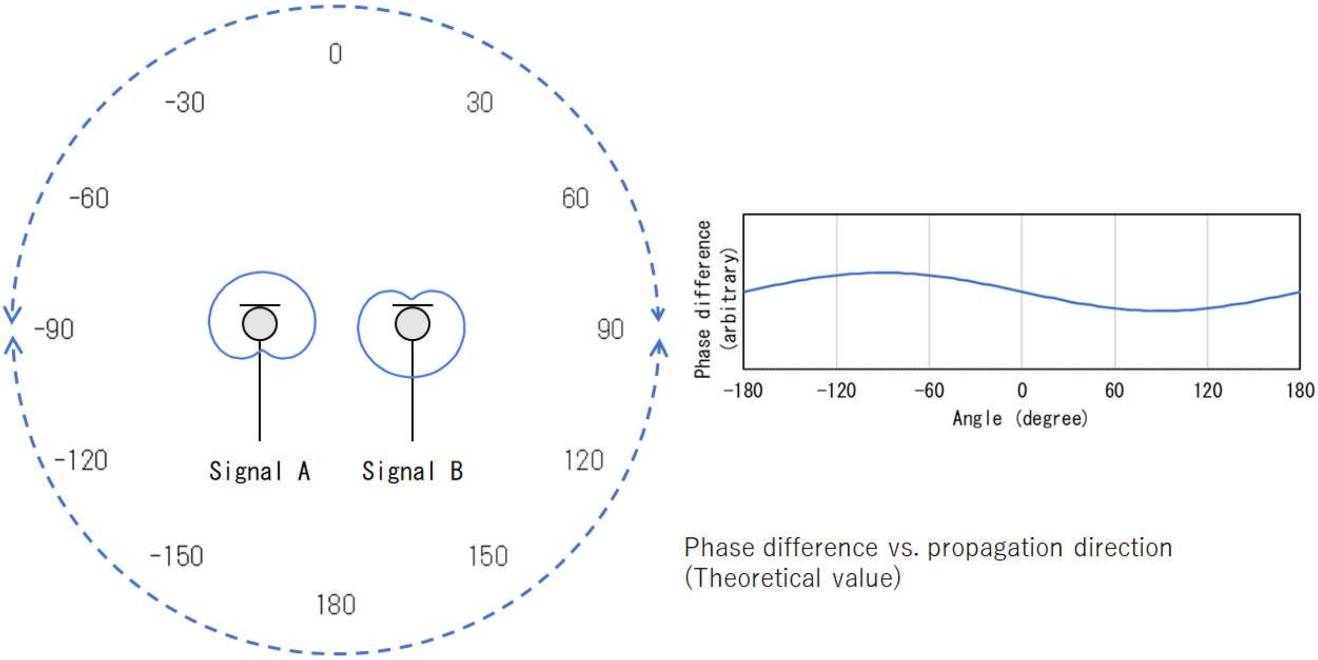

Fig. 8 Inter-channel phase difference
A top view of the system is shown on the left. Inter-channel phase difference (theoretical figure) is plotted on the right as a function of direction.

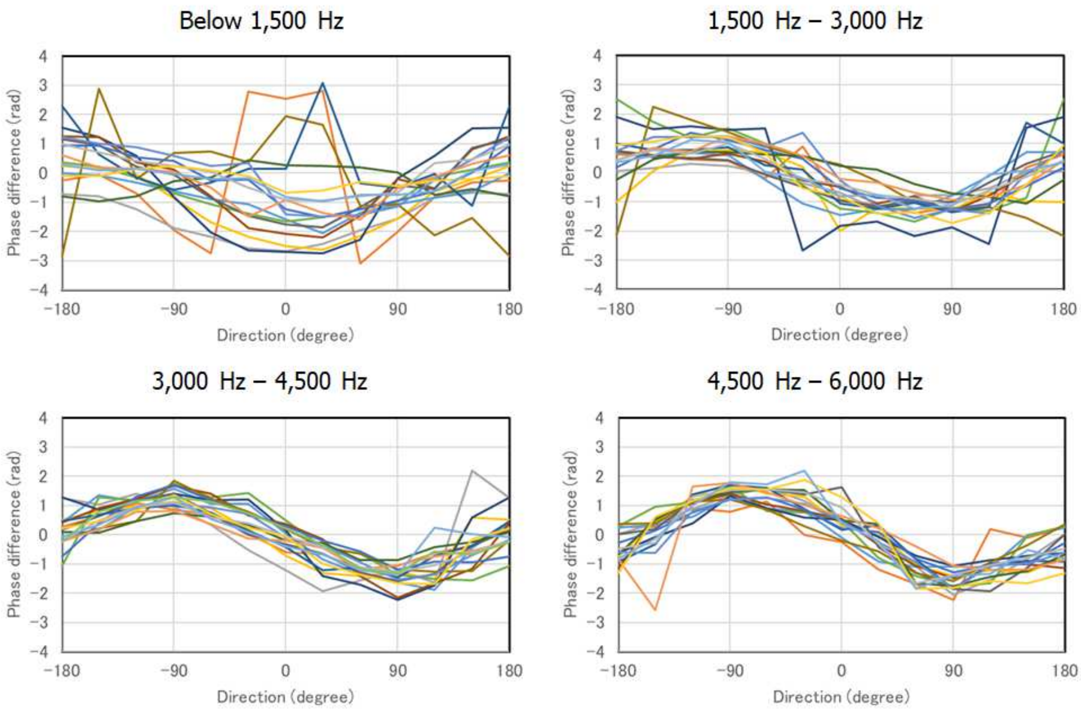

Fig. 9 Inter-channel phase difference
Inter-channel phase difference of TASCAM DR-22WL is plotted as a function of direction. Each line corresponds to each frequency bin.

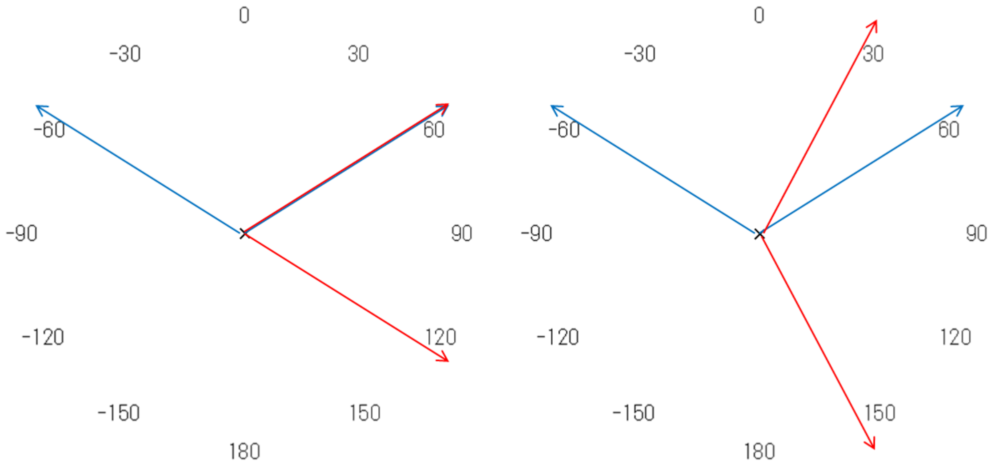

Fig. 10 Estimated propagation directions
Propagation directions estimated by the inter-channel level difference are represented by blue lines while those estimated by the inter-channel phase difference are represented by red lines. Two arrows (blue and red) point at the same direction in the left figure while they do not point at the same direction and the red one is closer to the median line in the right figure.

the graph in Fig. 8 in the higher frequency bands. It is indicated that the inter-channel phase difference (in the case of TASCAM DR-22WL) can be used for estimating the left-to-right lateralization of the sound at least at frequencies above about 1,500 Hz.

One-dimensional (front-to-rear) spatial information can be obtained from the inter-channel level difference and another one-dimensional (left-to-right) spatial information can be obtained from the inter-channel phase difference. By combining them, 2-dimensional spatial properties can be acquired.

It was confirmed by the measurement, that the propagation direction in a horizontal plane can be estimated for each frequency bin at least above about 1,500 Hz in the case of a handy recorder 'TASCAM DR-22WL' even when there is some reverberation.

3 Speed of propagation

In the section above, it was shown that the propagation direction at the observation point can be estimated for each frequency bin by using the inter-channel level difference and phase difference except for frequency bands below about 1,500 Hz. The other spatial property to be measured is the propagation speed. Theoretically, the faster the propagation speed is, the shallower the polar pattern of a cardioid microphone gets. In the proposed system, therefore, the faster propagation speed causes reduction of the inter-channel level difference. As the inter-channel level difference is reduced, the propagation direction estimated from the inter-channel level difference gets closer to the direction of either $\pm 90°$ than the actual propagation direction does.

The faster propagation speed results in the reduction of the inter-channel phase difference, too and the propagation direction estimated from the inter-channel phase difference gets closer to the median line ($0°$ or $180°$) than the actual propagation direction does.

Logically, we can speculate as follows.

As can be seen in the left illustration of Fig. 10, when the propagation direction estimated from the inter-channel level difference (blue arrow) overlaps the propagation

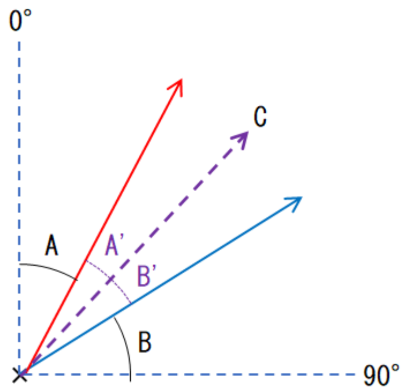

Fig. 11 Propagation direction and speed
When direction C is provided in such a way that the ratio between angles A and B equals that between angles A' and B,' C is supposed to be the actual propagation direction. When propagation speed is infinity, the sum of angles A' and B' is $\pi/2$.

direction estimated from the inter-channel phase difference (red arrow), they are supposed to indicate the actual propagation direction and the propagation speed in this case equals the sound speed. This is the case when there is only one sound source.

When the propagation directions estimated from the inter-channel level difference and phase difference do not overlap each other and the red arrow is closer to the median line as in the right illustration of Fig. 10, the propagation direction and speed in this case can be still estimated mathematically. When the ratio of angles A and B in Fig. 11 equals that of angles A' and B,' direction C is supposed to be the actual propagation direction. The faster the propagation speed is, the larger the angle between the arrows becomes. If the propagation speed equals the sound speed, two arrows overlaps each other pointing at direction C. When the propagation speed is infinity, the sum of angles A' and B' is $\pi/2$.

Assuming that C is the propagation direction, the inter-channel phase difference depends exclusively on the propagation speed. The direction of the red arrow, therefore, depends exclusively on the propagation speed. This means that the inter-channel phase difference can be calculated backward from direction C and the direction of the red arrow. Once the inter-channel phase difference is known, the propagation speed can be estimated from the propagation direction and the inter-channel phase difference.

4 Binaural reproduction

As described above, spatial properties of sounds at the observation point can be obtained by using only 2 microphones. By convolving the estimated spatial properties with each frequency component of the monaural signal at the observation point, binaural signals can be generated. This monaural signal has to be captured by an omnidirectional microphone placed at the observation point. The proposed system, however, does not have an omnidirectional microhone. In the proposed system, therefore, a sum of Signal A and Signal B was used, instead. Since their polar patterns are complementary each other, a sum of Signal A and Signal B gives an almost omnidirectional polar pattern.

5 Conclusion

In the proposed system, one-dimensional (front-to-rear) spatial information can be extracted from the inter-channel level difference and another one-dimensional (left-to-right) spatial information can be extracted from the inter-channel phase difference. Accordingly, 2-dimensional spatial sound can be conveyed by a simple stereo audio data stream.

Since some handy recorders such as ZOOM H6 and TASCAM DR-22WL already have the microhone system that satisfies the requirement for capturing 2-dimensional panoramic sound, they can be used for this purpose and no additional hardware is needed.

When using TASCAM DR-22WL, the inter-channel phase difference at the frequencies below about 1,500 Hz is not supposed to be reliable. Further improvement is required.