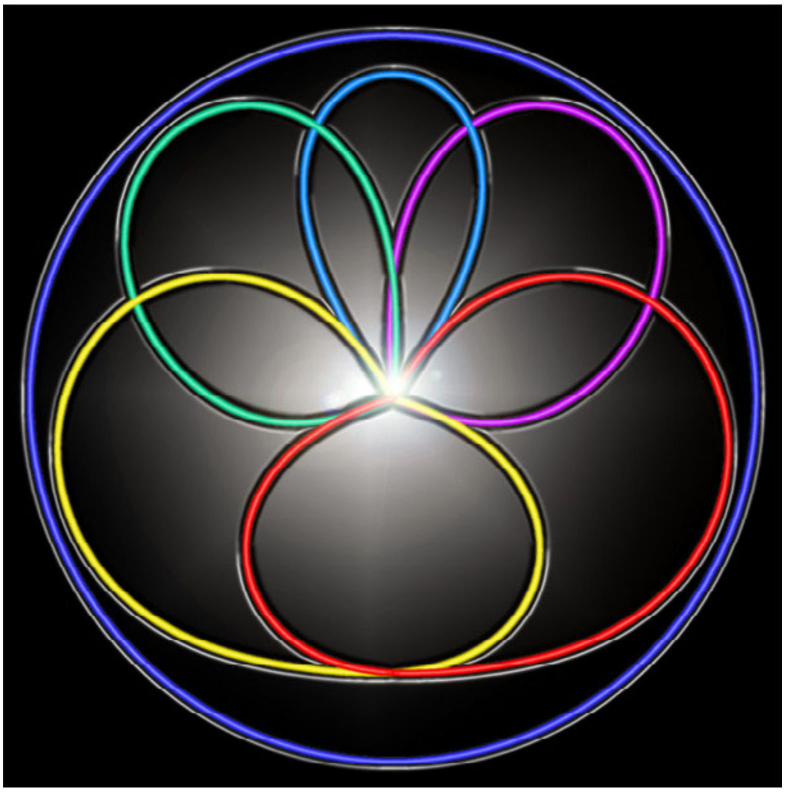
5.0 Sound recording

in

High Spatial Resolution



Trinnov Audio – 30-32 avenue de la République – 94815 Villejuif Cedex – France
Tel : +33 6 16 81 62 87 – Fax : +33 1 46 81 77 27
www.trinnov.com – contact@trinnov.com

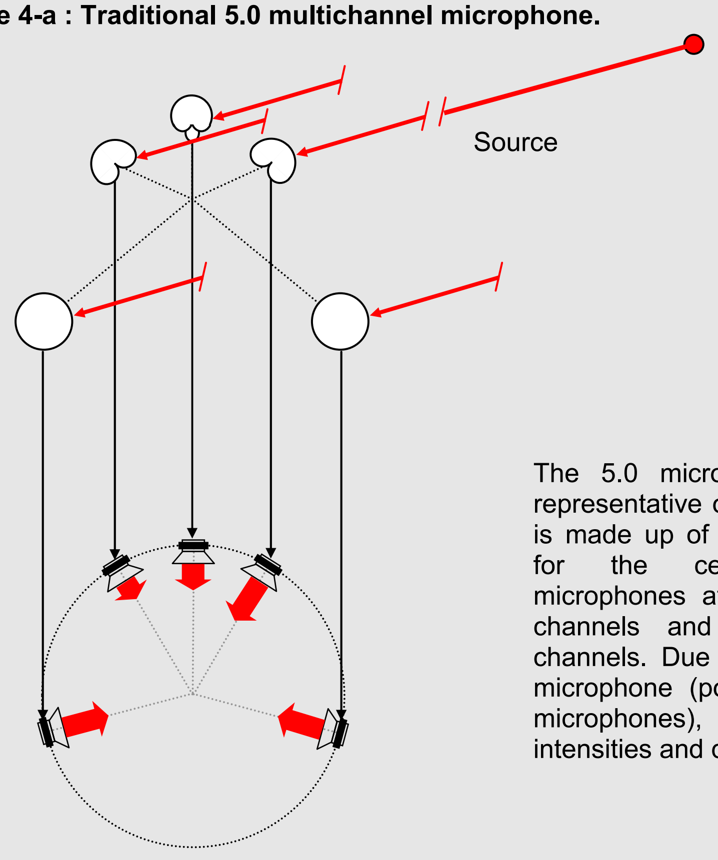
5.0 Sound recording in High Spatial Resolution

1 High Spatial Resolution

1.1 The limits of traditional multichannel sound recording

Currently all multichannel recording systems attempt to use the directivities of current microphones as much as possible. The approach used is directly based on stereophony and consists in producing differences in intensity and delay between the signals captured by acting on microphone orientations and their respective distances. The position of sound sources surrounding the microphone is "recorded" thanks to the differences in intensity (ΔI) and time (ΔT). With this approach, the difficulty resides in determining the optimum combination of orientation and distance between microphones. This approach is shown in figure 4-a.

Figure 4-a : Traditional 5.0 multichannel microphone.



The 5.0 microphone shown opposite is representative of the traditional approach. It is made up of a hypercardioid microphone for the central channel, cardioid microphones at 45° for the right and left channels and omnidirectional for back channels. Due to the geometry of the 5.0 microphone (position and direction of the microphones), the channels have different intensities and delays.

And yet this approach is fundamentally limited as current microphones only have a restricted number of directivities. This range of directivities goes from omnidirectional to bidirectional (figure-of-8), including intermediate directivities (infracardioid, cardioid, hypercardioid). In this way, traditional sound recording systems are unable to reproduce the entire range of signals which multichannel broadcasting systems can reproduce. It would also seem appropriate to consider whether this approach actually provides optimum multichannel sound recording.

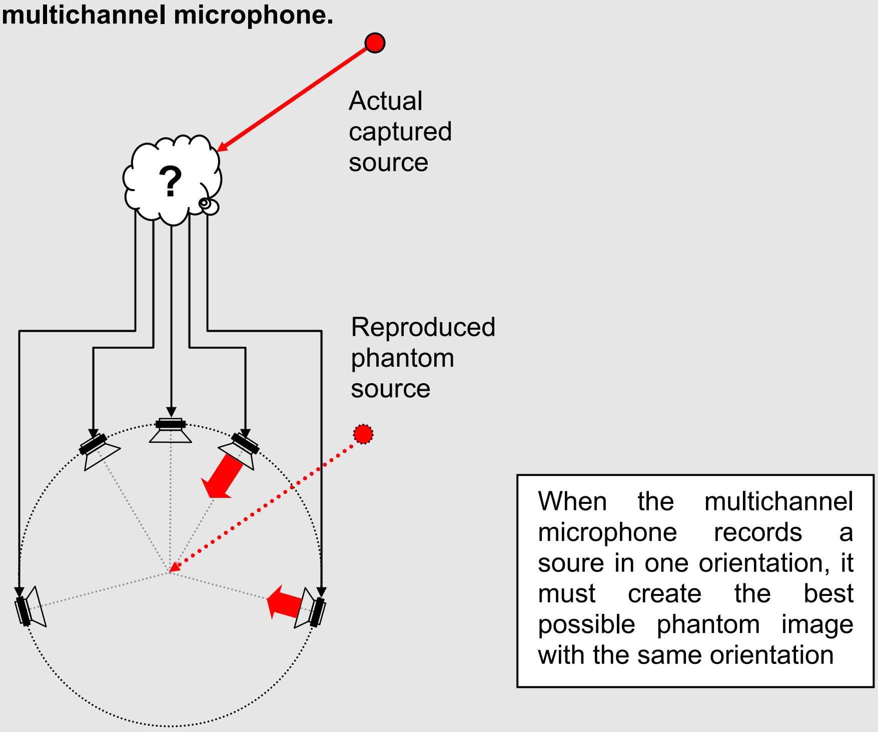
In addition, although fundamentally limited, the traditional approach allows for the definition of a large number of multichannel microphones whose differences would be clearly audible. Simply by using 5 types of directivity, 10 orientations and 50 positions (5 angles and 10 radiuses), over 15 thousand million different recording systems can be produced! It is therefore unrealistic to hope to list and test all systems. We point out that, with the same possibilities, only 2500 stereophonic systems exist. Consequently the switch from 2 channel stereophony to multichannel (5 channels) leads to a multiplication of possible combinations in terms of sound recording systems. From this point of view, the empirical approach of testing/errors on which stereophony is based also appears as limited.

1.2 What is optimum multichannel sound recording

Faced with the complexity of multichannel sound recording, **Trinnov** *Audio* tackles the problem from an entirely new angle. To begin with, let us forget the directivities of current microphones and consider the following question: **what would be the ideal behaviour of a system allowing for optimum 5.0 sound recording over 360°?**

The answer is trivial: When seeking the ideal recording of a sound source in a given orientation, the 5.0 restitution system must provide the best phantom source possible with the same orientation as the recorded source. Ideally speaking, the only limitation tolerated is the intrinsic limitation of the 5.0 restitution system. The sound recording system must not provide additional limits. The concept of an ideal multichannel microphone is shown in figure 4-b.

Figure 4-b : Ideal 5.0 multichannel microphone.



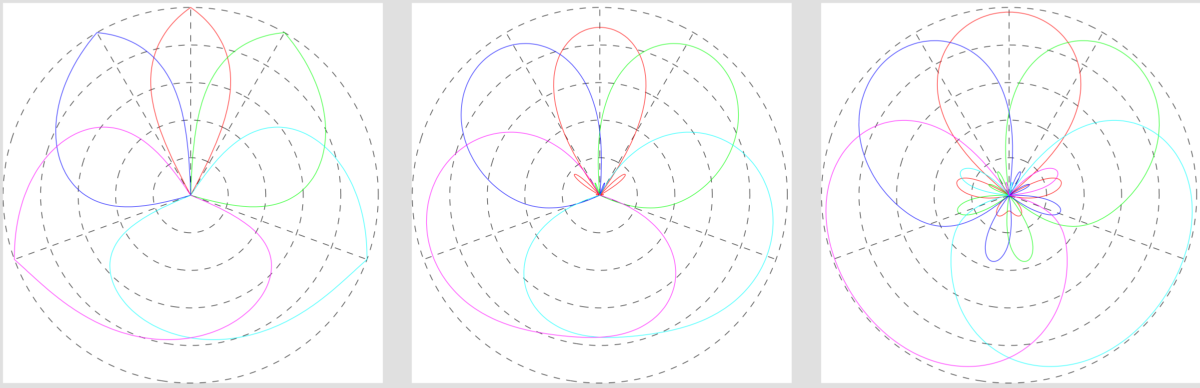
Defined in this way, optimum multichannel sound recording corresponds to panning laws (also known as pan-pot laws or directed monophony laws). These laws have been precisely studied to form optimum phantom images over 360°. A 5.0 panning law is defined by 5 directivity patterns showing the intensity of channels and forming an optimum phantom image in each orientation. **In this way, the directivities of a panning law may be interpreted as ideal directivities for optimum multichannel sound recording.** Many panning laws exist, but with a few differences, these laws tend to have the same general shape. To give an example, here are 3 panning laws:

- Intensity panning
- Order 5 intensity panning. The concept of "order" will be defined later.
- Continuous surround panning law, order 4 (P. Craven, 2003).

These 3 examples are shown below.

Figure 4-c: Ideal directivities for 5.0 optimum sound recording

Intensity panning. | Order 5 intensity panning. | Order 4 continuous surround.



With just a few differences, these laws all have the same general shape.

1.3 Concept of high spatial resolution multichannel sound recording

The concept of high spatial resolution is directly inspired by the concept of high temporal resolution (also known as high fidelity).

High temporal resolution refers to when a signal is recorded very precisely even with rapid variations. To be able to follow rapid variations, the recording system must be able to record high frequencies and must therefore have an extended frequency band (up to at least 20kHz). Consequently, high temporal resolution requires a wide frequency band.

In parallel, high spatial resolution refers to when the acoustic field is recorded very precisely even with rapid spatial variations. Spatial variations correspond to the angular distribution of the acoustic field around the microphone. To be able to follow rapid spatial variations, the recording system must be able to record high spatial frequencies and must therefore have a wide spatial frequency band.

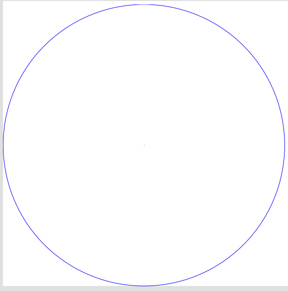
A spatial frequency actually corresponds to a directivity pattern for which the precision of lobes directly determines spatial precision. The more the directivity pattern is selective, the higher the spatial resolution. A selective directivity pattern is more able to determine sound sources in space. In order to define a spatial spectrum, spatial frequencies are classified depending on their spatial precision or "order". Figure 4-d shows certain spatial frequencies up to order 5.

As already highlighted, current microphones only have a restricted number of directivities. In reality, these directivities are only mixes of omnidirectional and bidirectional directivity. Consequently, possible spatial resolution is systematically limited to order 1.

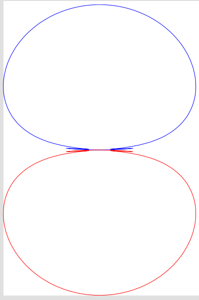
On the contrary, ideal directivities for optimum 5.0 sound recording correspond to a mix of spatial frequencies reaching order 5. In this way, optimum 5.0 sound recording requires spatial precision of order 5. In other words, optimum spatial precision is 5 times higher than that obtained with current microphones. The concept of high spatial resolution is therefore based on advanced scientific and technical concepts. Figure 4-e shows the difference in selectivity between order 1 and order 5 for hypercardioid and cardioid directivities.

Figure 4-d: Spatial frequencies of up to order 5.

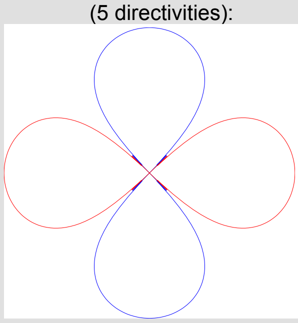
Spatial frequency order 0
(omnidirectional directivity)



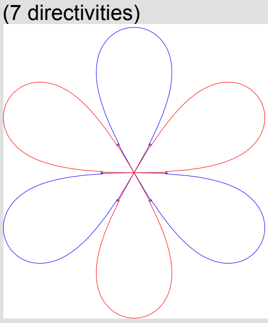
Spatial frequency order 1
(3 bidirectional directivities)



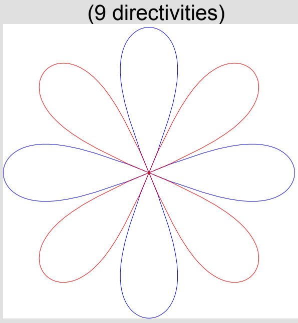
Spatial frequency order 2
(5 directivities):



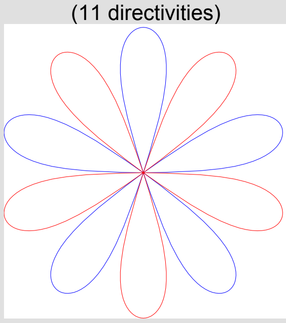
Spatial frequency order 3
(7 directivities)



Spatial frequency order 4
(9 directivities)



Spatial frequency order 5
(11 directivities)

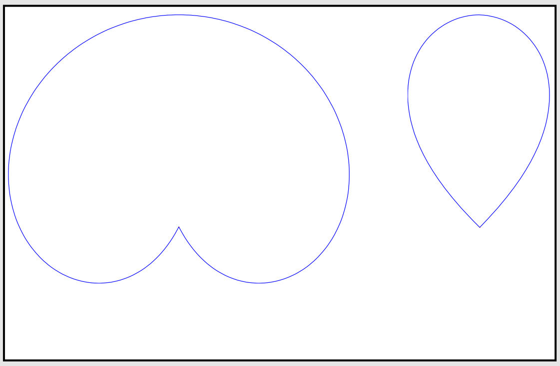


With these figures, positive spatial frequency lobes are in blue and negative lobes in red.

Figure 4-e: Comparison of spatial resolution order 1 and order 5.

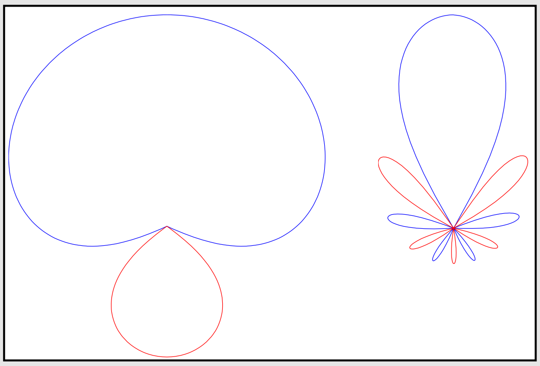
Cardioid:

Order 1 Order 5



Hypercardioid:

Order 1 Order 5



With these figures, positive spatial frequency lobes are in blue and negative lobes in red.

If highly selective microphones are available, high spatial resolution sound recording can be obtained very simply. All that is needed is 5 coincident microphones whose directivities correspond to the ideal directivities defined previously.

1.4 High spatial resolution requires a new approach

The directivities required to obtain high spatial resolution 5.0 sound recording present the following characteristics:

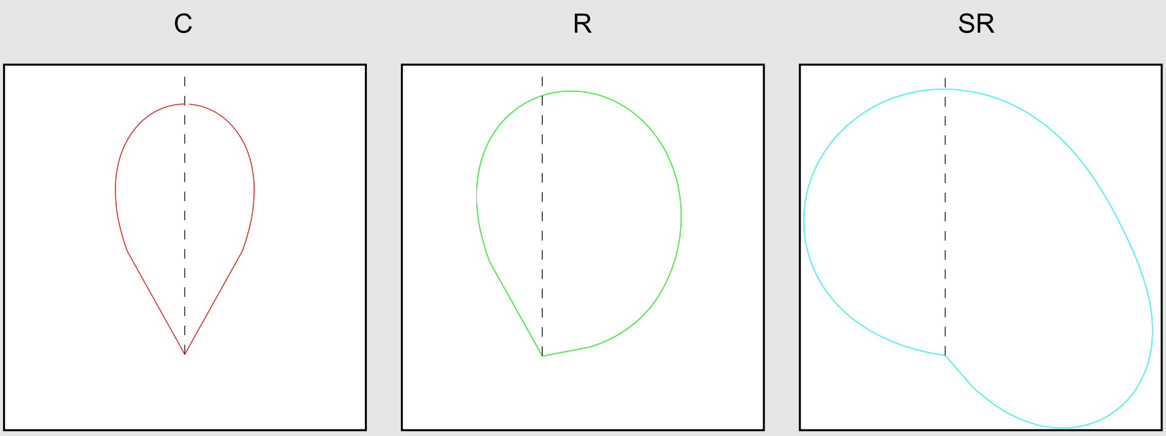
- Highly selective. This spatial selectivity is at the origin of high spatial resolution. This is obvious for the 3 frontal directivities which provide order 5 selectivity for the entire frontal zone (+30° / -30°). As already mentioned, the selectivity required is 5 times higher than the selectivity of current microphones. We also note that back directivities also have high spatial selectivity (order 3 spatial frequency band).

- Asymmetry. With the exception of the directivity of the central channel, the other 4 directivities are asymmetrical. This characteristic is indispensable if we require optimum adaptation to the irregularity of 5.0 configuration. If the loudspeakers had been placed in a regular pentagon (0°, +/-72° and +/-144°) ideal directivities would have been symmetrical. Current microphones are unable to produce asymmetrical directivities.

- Complementary. Directivities which produce phantom images between 2 loudspeakers are entirely complementary to ensure that phantom images are produced with the same intensity in the zone between the two loudspeakers.

These characteristics are shown in figure 4-f.

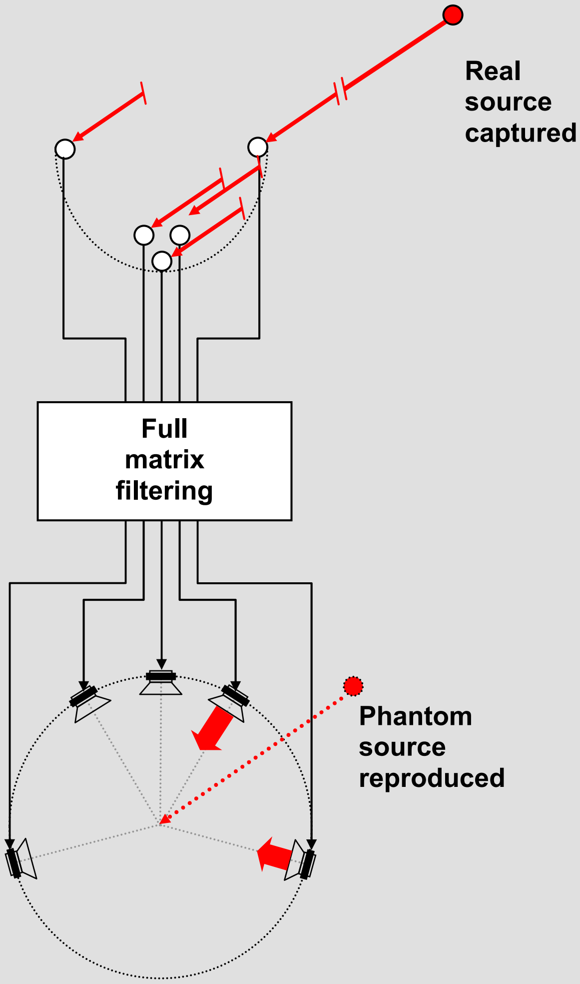
Figure 4-f: Selectivity, asymmetry and complementarity of ideal directivities.



The directivities required for recordings with high spatial resolution are not obtained naturally by current microphones. To obtain required performances, it is necessary to **abandon the traditional approach of capturing each channel with one capsule. A more complete approach based on matrixing must be considered.**

In order to obtain maximum spatial quality, **Trinnov** *Audio* has developed the ultimate solution in terms of linear processing: full matrix filtering. Each channel is obtained by taking advantage of the entire microphone array. The signal from each microphone is filtered in a specific manner. The group of signals thus formed is combined in order to produce one optimum channel. Only an advanced system of this type can provide high spatial resolution. Figure 4-g shows a multichannel microphone with high spatial resolution based on this layout.

Figure 4-f: 5.0 multichannel microphone based on full matrixing.



2 Advantages of high spatial resolution

As compared with a multichannel recording at standard resolution, a high spatial resolution recording provides 5 times better selectivity, particularly with the frontal zone. With a high spatial resolution recording, only the two loudspeakers immediately next to each phantom source are active. On the contrary, with standard resolution recording, all loudspeakers are active. This causes complex interactions which worsen the spatial quality of the recording. Figure 4-g compares high spatial resolution recordings with standard spatial resolution recordings.

High spatial resolution sound recording has many advantages as compared with standard resolution sound recording. Particularly:

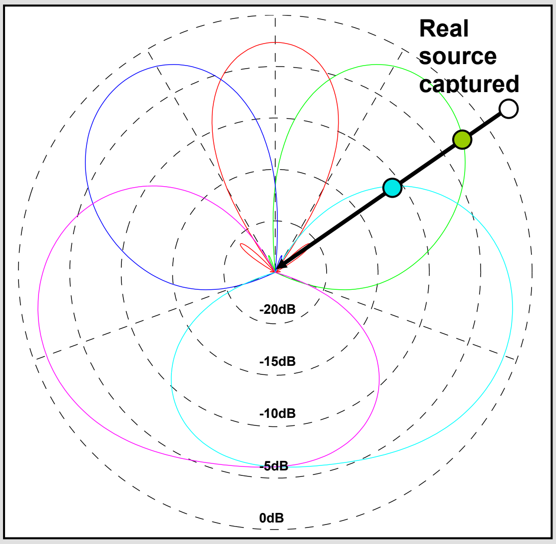
- Optimising channel **separation**.
- Increasing source **precision** (punctuality).
- Increasing the listening zone (**sweet spot)**.
- Respecting **tones** independently of source orientation.
- Producing phantom images over **360°,**
- Allowing for easy and good quality **down-mix**, without modifying tone or distorting angles.
- Providing total **compatibility** with usual panning laws, making adding spot microphones easier.

- Controlling source spatialisation outside the azimuthal zone.
- Allowing better control of **distance factors.**
- Providing better control for the coverage angle thanks to precise control of the orientation of reproduced images.
- Providing better robustness for **processing**: treating a channel does not affect phantom images other than in the surrounding area.

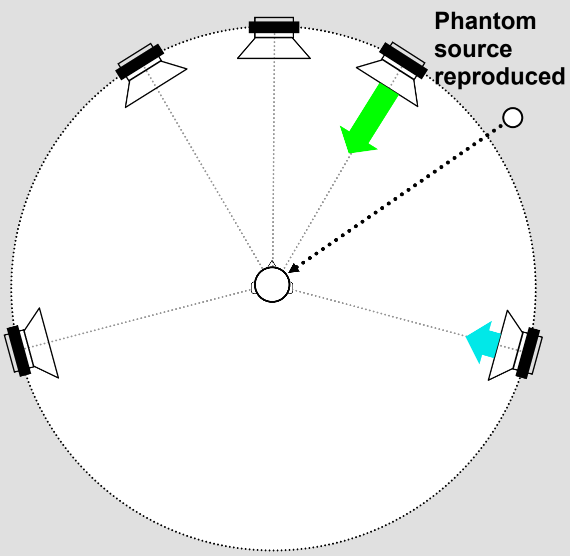
Figure 4-g: Comparison of 5.0 high spatial resolution sound recording with standard spatial resolution recording.

High spatial resolution (order 5)

Capturing directivity

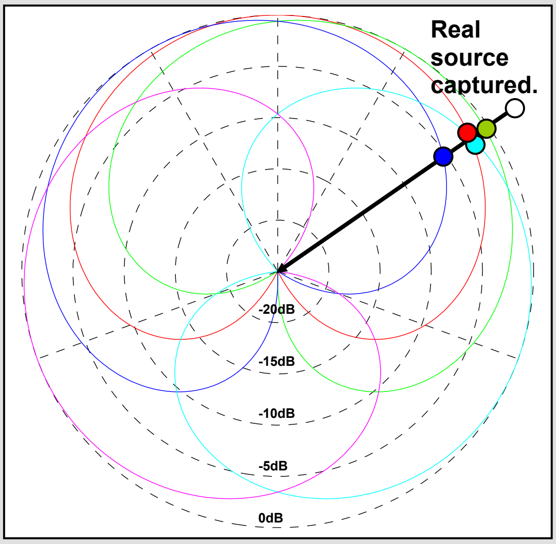


Restitution on 5.0 configuration
(standard ITU-R BS 775-1)

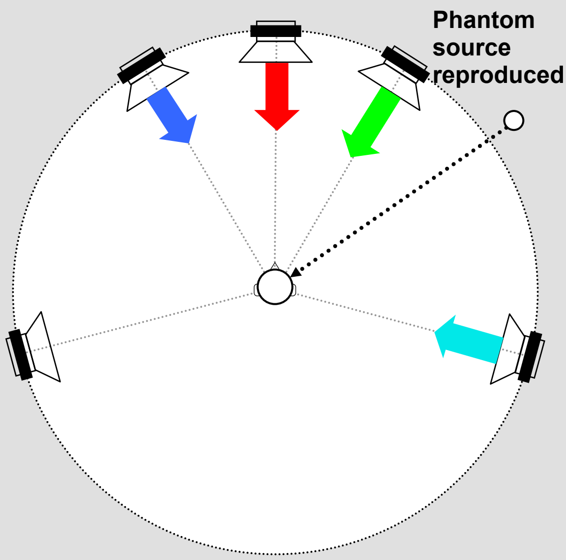


Standard spatial resolution (order 1)

Capturing directivity
(5 cardioid microphones oriented according to ITU format directions)



Restitution on 5.0 configuration
(standard ITU-R BS 775-1)



The advantages of high spatial resolution sound recording are presented in more detail in each of the following sections.

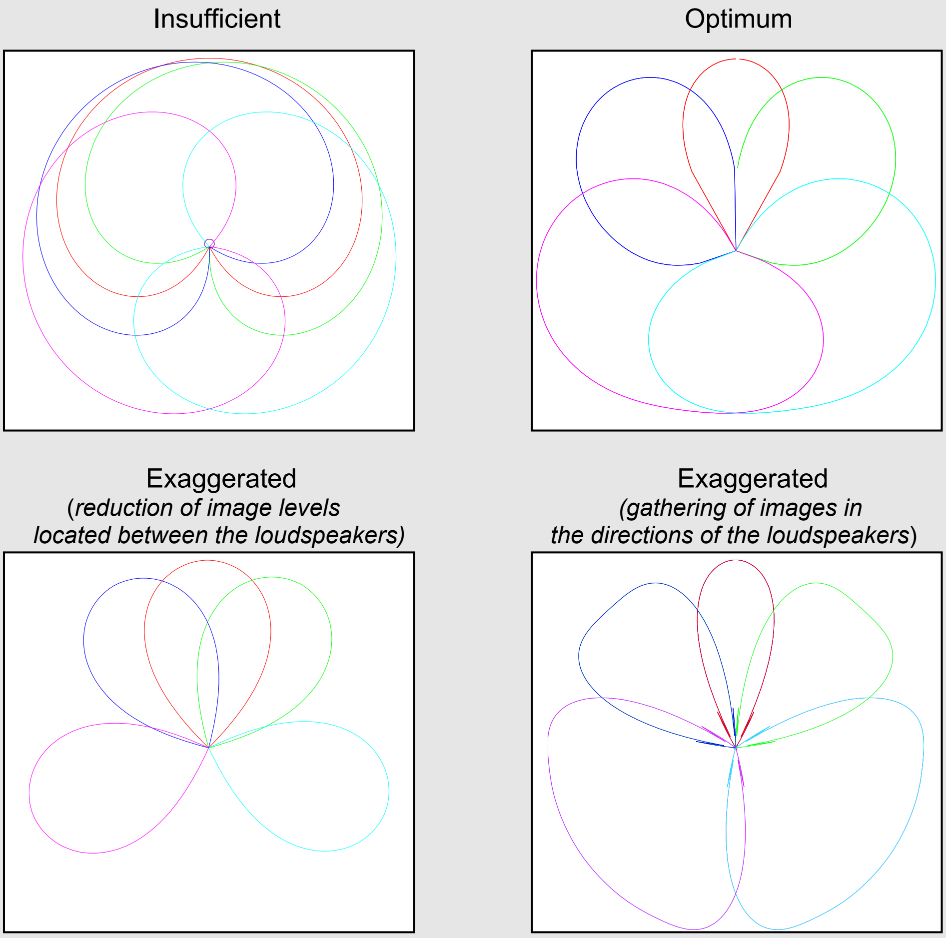
2.1 Optimum channel separation

With a multichannel signal, global spatial quality is directly linked to channel separation (or decorrelation) **Insufficient separation leads to limited spatial performances.** When separation increases, overall spatial quality also increases. However, a limit exists, beyond which any increase in separation reduces performance. This correlation is indispensable for producing phantom images. If too much separation exists, two undesirable phenomenons appear:

- The phantom images located between the loudspeakers reduce sound levels and "holes" form in the sound envelopment.
- Phantom images appearing between the loudspeakers move towards the nearest loudspeakers. The images gather together at the loudspeakers and "holes" form in the sound envelopment.

Should no correlation exist, the 5 channels will carry totally different signals and the sound space produced will correspond to 5 sources in the directions of the 5 loudspeakers. Therefore, should no correlation exist, no phantom images exist. **In this way, exaggerated separation deteriorates phantom images.** Consequently, an optimum level of channel separation exists. **The aim of the high spatial resolution 5.0 microphone is therefore to achieve optimum separation** (and not maximum) for channels. Figure 4-h shows the concept of optimum channel separation.

Figure 4-h: Channel separation.

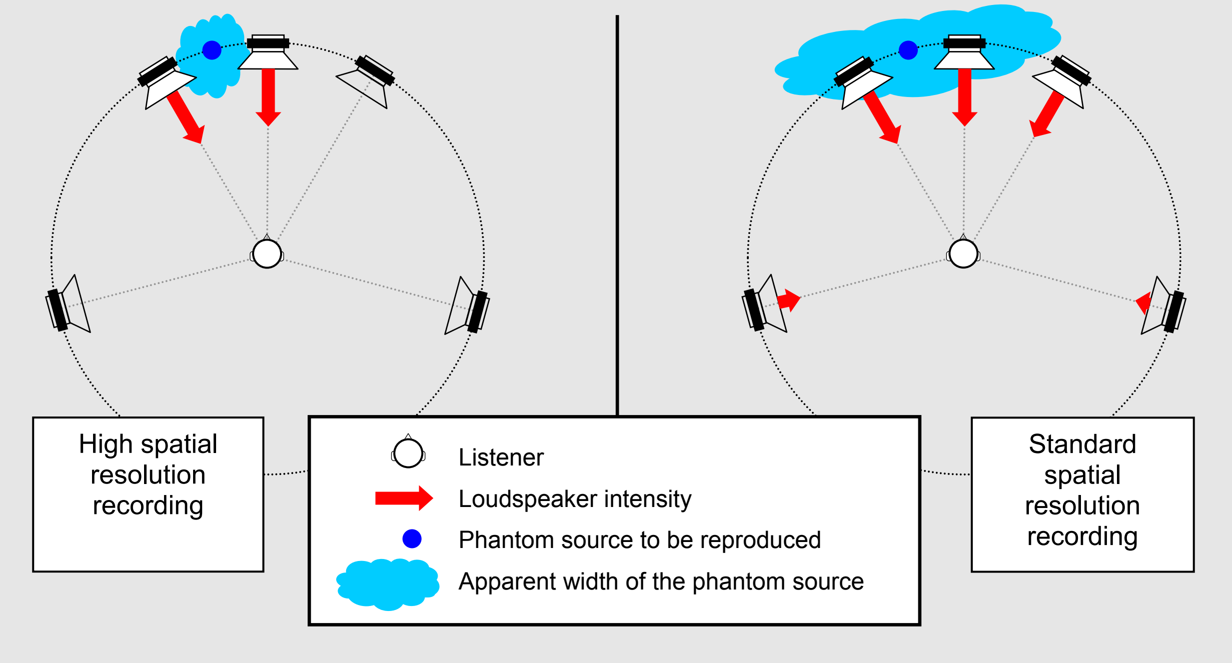


2.2 Increasing source punctuality

With standard resolution recording, the signal emitted by a source is reverberated on almost all channels. Consequently, the corresponding phantom image is reproduced by a high number of loudspeakers. In reality, the contribution of a high number of channels in reproducing a phantom source is generally a problem. To give an example, when 5 channels are active, they simultaneously reproduce 10 phantom images. A group of 5 loudspeakers allows for the definition of 10 couples which each produce one phantom image. Therefore, multichannel sound cannot be reduced to a combination of 5 images appearing between the combined pairs of loudspeakers (e.g. between the C and R channels). The 5 other "static" phantom images forming between the combined pairs of loudspeakers (e.g. between the C and R channels) must also be considered. And yet, these 10 phantom images appear in different directions and are at the origin of contradictory spatial information. The result is spatial distribution of the phantom sources. Conflicts between phantom images are problems which are specific to multichannel systems and do not appear in stereophony as only 2 channels and 1 phantom image exist.

With high spatial resolution recording, only 2 channels mainly contribute to the reproduction of the phantom image. Therefore only one single phantom image exists and no contradiction will occur. The reproduced source is well situated and highly punctual.

Figure 4-i: Increased source punctuality.



We note that high spatial resolution accepts more than two active loudspeakers to produce precise phantom images. In this case, the system is optimised to ensure that the overall contribution of the 5 channels leads to the precise reconstruction of the different spatial frequencies in the acoustic field in the sweet spot.

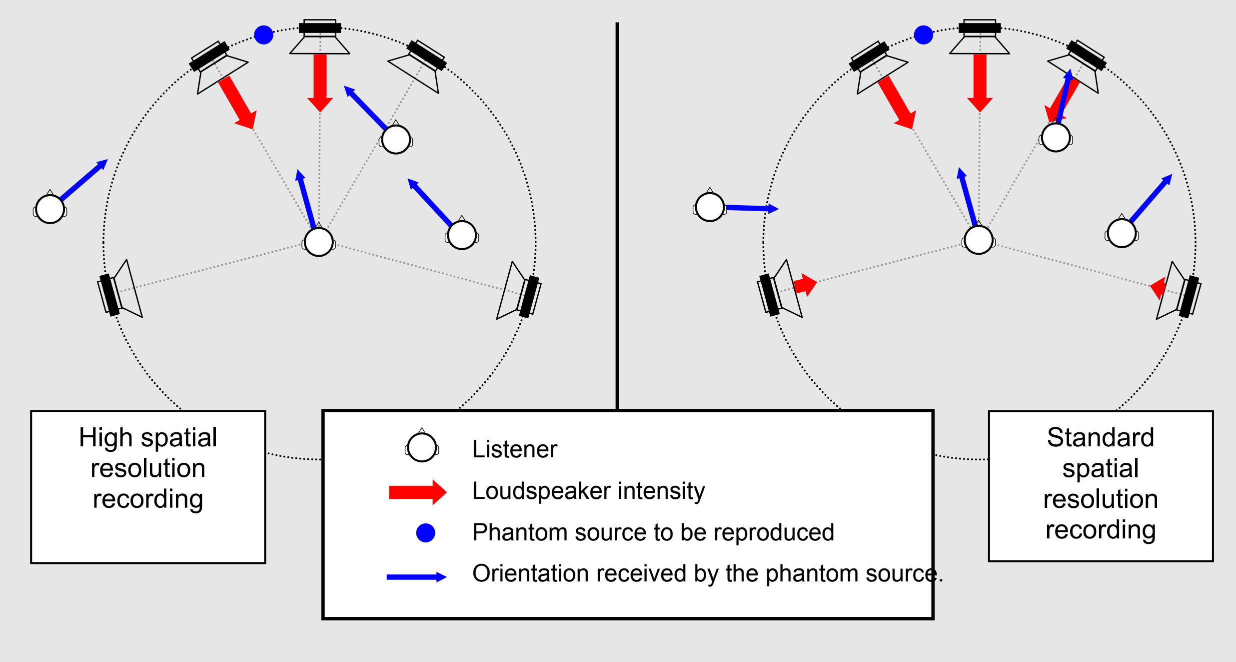
2.3 Increasing the size of the sweet spot.

With standard spatial resolution recording, the signal emitted by a source is passed on almost all channels. When the listener moves in the room, he or she will note a dominant intensity coming from the nearest loudspeaker. The reproduced phantom source is attracted in the direction of the dominant loudspeaker. This leads to visible instability of phantom sources with the movements of the listener. The zone for which spatialisation is suitably maintained is not wide.

On the contrary**, with a high spatial resolution recording, only the loudspeakers immediately next to the phantom source are active**. Independently of the position of the listener, the phantom image remains located between the two nearest loudspeakers. This leads to substantial stability of phantom sources with the movements of the listener. Obviously, stability is also linked to the difference between the loudspeakers and the most stable images will be obtained for the frontal zone (+30°/-30°). We reiterate that the ITU 5.0 configuration has been specifically studied to produce a highly stable frontal sound zone. **Only a high spatial resolution microphone can use this configuration in an optimum manner to provide a maximum sweet spot.**

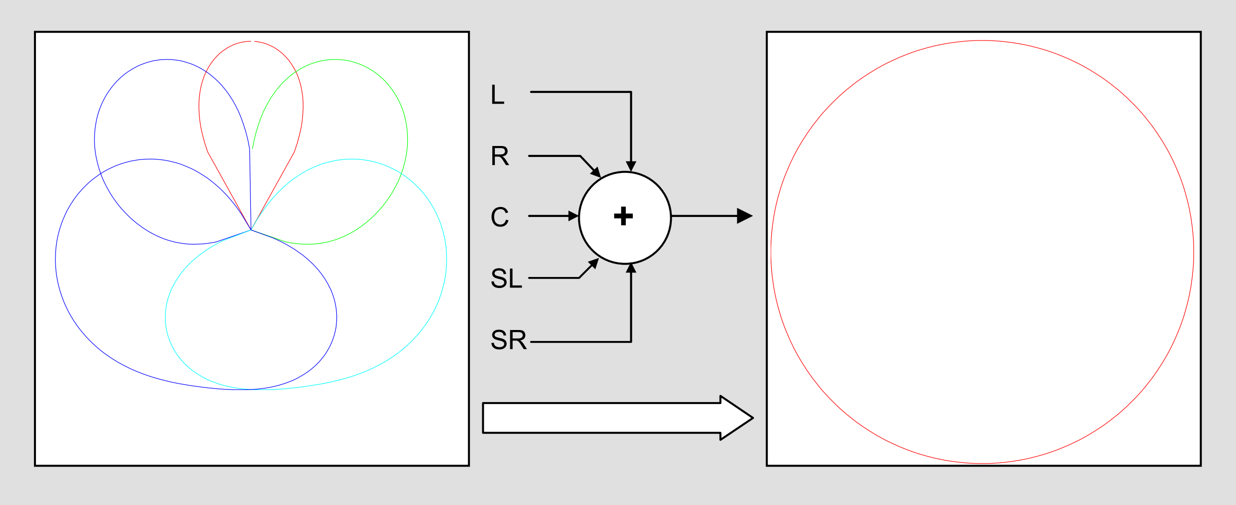
Figure 4-j shows the orientation received in the different listening positions.

Figure 4-j: Increased source punctuality.



2.4 Maximum respect of tones in all directions

Figure 4-k: Respect of tones for all directions



The respect of tones is a guaranteed characteristic of high spatial resolution recording. A multichannel sound recording system respects tones if an omnidirectional microphone placed at the centre of an ideal restitution system delivers the same signal as an omnidirectional microphone placed at the recording point. However, at the centre of the restitution system, all 5 channels combine perfectly. Therefore, tones are respected if, by combining the 5 channels of 5.0 signal, we find the exact same signal which would have been captured by an omnidirectional microphone. In other words, **5.0 sound recording maintains tones if the down-mix to mono format is perfect** (see section 2.6). In more scientific terms, this means that the **order 0 spatial frequency of the acoustic field (omnidirectional component) must absolutely be captured and reproduced with extreme precision for all frequencies.** This characteristic is ensured by high spatial resolution microphones and is shown in figure 4-k.

Spatial resolution may also vary with frequency, without affecting tone. This is a new outcome provided by high spatial resolution and which is worth explaining. As a general rule, a microphone whose directivity varies with frequency inevitably introduces tone modifications. For a given direction, a microphone captures sound with a sensitivity which varies depending on frequency, thereby introducing changes in tone. This problem cannot be solved when considering one single microphone. On the contrary, with a multichannel system, it is indispensable to consider all channels simultaneously. And yet, different resolution systems are perfectly able to form phantom images whose orientation and tone are very precisely reproduced. In this case, the only differences appear in spatialisation. It is therefore entirely possible to imagine an extremely advanced system with high resolution in the band 300Hz – 5000Hz and a resolution which reduces gradually at the ends of the frequency band (20Hz – 20000Hz). To confirm our theory, we reiterate that this concept is already used as a general rule with 5.1 format. In this format, very low frequencies (< 80Hz) are transmitted by a single "sub" channel and do not include spatialisation. Despite a brutal transition in the spatial resolution, tones are suitably respected on the entire frequency band, including for the lowest frequencies. Obviously, the same can be said for slower transitions between resolutions. Consequently, a system whose resolution varies with the frequency may be envisaged and lead to high performances without changing the tone of phantom sources on the entire frequency band. Overall spatial performances depend extensively on the spatial resolution achieved for each frequency. This must absolutely be high in the band 300 Hz – 5000 Hz.

2.5 Presence of phantom images over 360°

High spatial resolution guarantees the presence of phantom images over 360°. This is the direct consequence of the **respect of order 0 spatial frequency (omnidirectional component) in the acoustic field.** All sound sources surrounding the microphone will be reproduced at the correct level and will form phantom images. However, this is insufficient to guarantee the control of the distribution of sources upon restitution.

2.6 Easy and good quality down-mix

Down-mix problems have been studied since the appearance of stereophonic format in order to ensure descending compatibility with monophonic format. For multichannel formats, the problem of compatibility becomes far more delicate as 5.0 format must be compatible with all formats, including the most frequent formats: 4.0 "Quad", 4.0 "Surround", 3.0 "Surround", 2.0 "Stereo" and 1.0 "Mono".

The ability of a sound recording system to accept down-mix is linked to the implementation principle.

- <u>With coincident sound recording</u>, optimum down-mix is easily obtained by recombining the different channels by very simple matrixing. Spatialisation is achieved uniquely via variations in intensity (ΔI) between channels. In this way, mixing channels only modifies spatialisation without affecting tone.

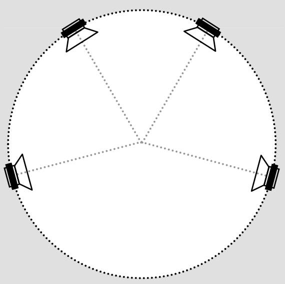
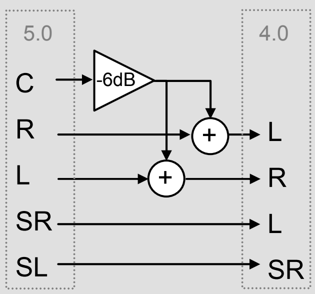
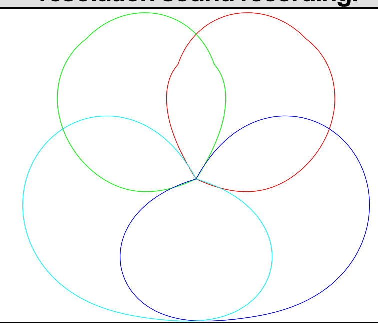
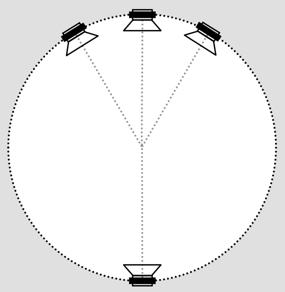
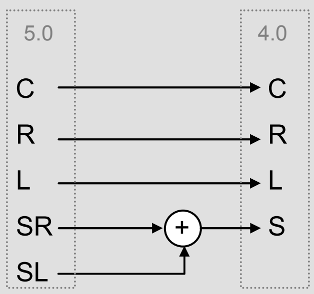
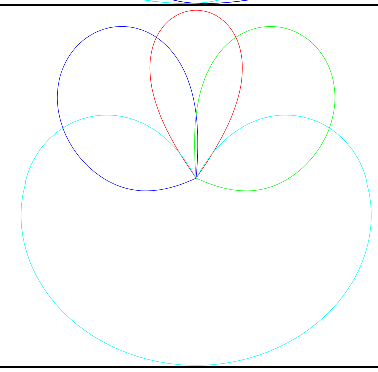
- <u>With non-coincident sound recording</u>, down-mix is extremely delicate and never produces optimum results. These systems introduce time differences (ΔT) between channels which makes mixing between them impossible. Simple matrixing of non-coinciding microphones systematically produces tone alternation (comb filter effect) and spatialisation alteration (spatial aliasing effect: appearance of several uncontrolled secondary lobes).

The signal provided by a high spatial resolution system is based on a coincidence principle. Spatialisation is achieved uniquely via variations in intensity (ΔI) between channels and no variation in time is introduced (ΔT). Consequently, **high spatial resolution recording allows for easy and high quality down-mix.**

As compared with standard resolution sound recording, high spatial resolution provides a decisive advantage: **the down-mix of a high spatial resolution sound recording retains the properties of a high spatial resolution sound recording. The recording resolution simply and automatically adapts to the maximum resolution accepted by the new format.** In this way, the down-mix of a high spatial resolution recording will remain similar to the results which would have been obtained by applying the principle of high spatial resolution directly to the new format. This is known as corresponding high spatial resolution sound recording. However, the performances achieved will be slightly less perfect due to lesser optimisation. The mono down-mix will include no spatialisation and the result obtained is identical to that obtained with monophonic sound recording.

Particularly, **the down-mix operation does not alter tone or distort angles in phantom images.** The phantom images formed by 4.0 and 3.0 formats are located in the same orientations as those of 5.0 format. Only performances in terms of spatialisation (source punctuality, sweet spot size, …) are reduced. With 2.0 format, the directions of phantom images are respected for all of the frontal zone (from +30° to –30°). Outside of the frontal zone, images are reflected onto the frontal zone while perfectly maintaining lateralisation and tone. Table 4-I shows the down-mix operation of a high spatial resolution sound recording for the main multichannel formats.

Figure 4-I: Down-mix of a high spatial resolution sound recording.

Format	Configuration	Down-mix matrix	Corresponding high spatial resolution sound recording.
4.0 "Quad" Spatial resolution: Order 3			
4.0 "Surround" Spatial resolution: Order 5			

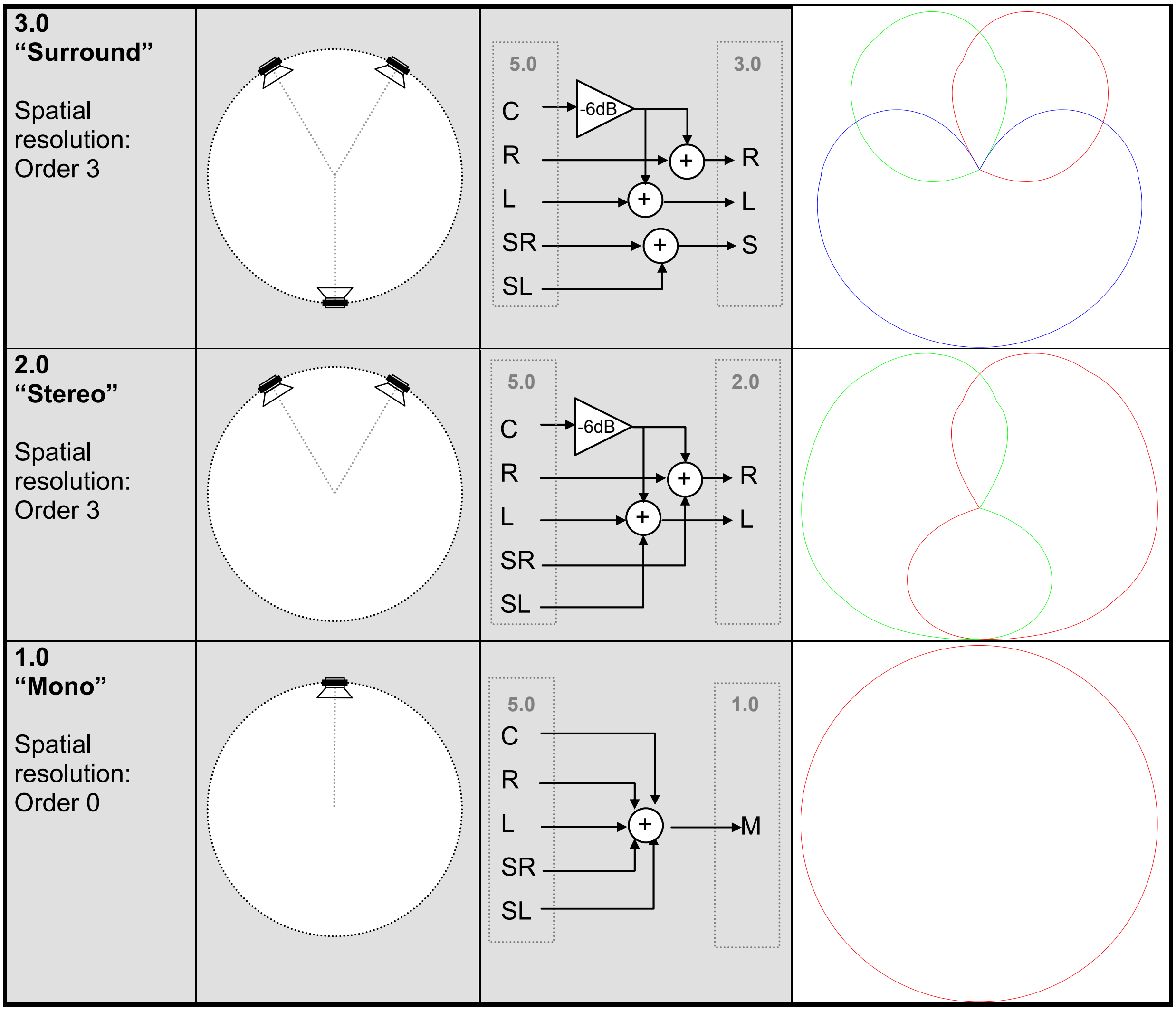
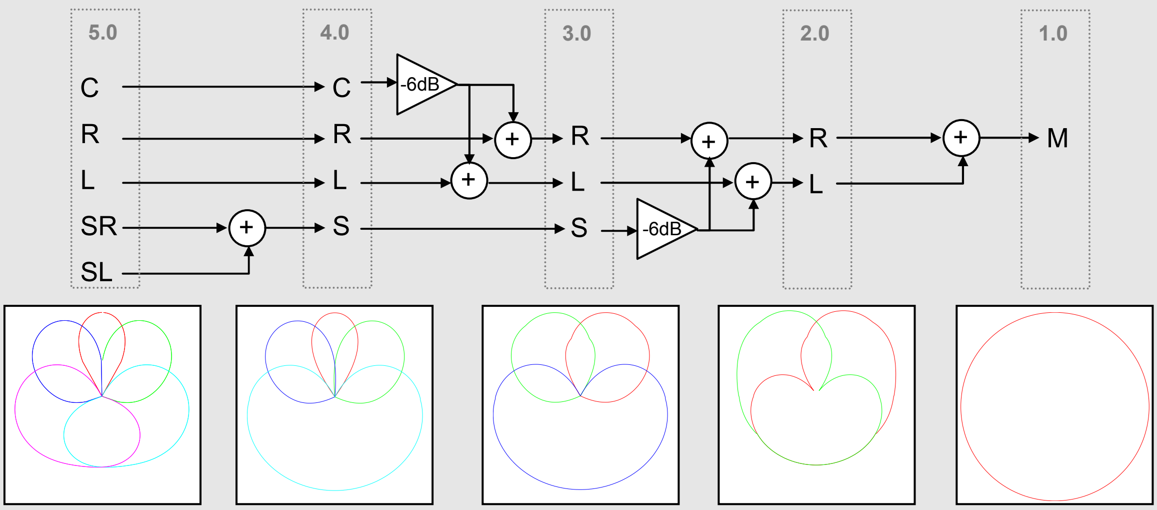


Figure 4-3m: Accumulation of down-mix towards hierarchically lower formats.



Subject to due compliance for format hierarchy, **high spatial resolution accepts accumulated down-mix via the different formats.** In this case, the final down-mix will be identical to a direct down-mix towards the final format and will therefore maintain all the qualities of high spatial resolution. To give an example, initial 5.0 format may be down-mixed in 4.0 Surround format, and then down-mixed in 3.0 Surround format, before being down-mixed in 2.0 Stereo format and finally down-mixed in mono format. The mono signal thus obtained will be identical to the signal obtained with a mono down-mix directly applied to the initial 5.0 format.

2.7 Total compatibility with regular pan-pot laws

The behaviour of a high spatial resolution microphone is similar to that of panning laws (see section 1.2). Consequently, **high spatial resolution sound recording mixes perfectly with artificially spatialised mono sounds using panoramic laws.** In both cases, the sound produced is similar and the mix is homogenous.

This is a new flexibility provided by high spatial resolution. Standard resolution sound recording is, by definition, restricted to order 1 precision, while a spatialised source with panoramic laws is high spatial resolution. Obvious differences appear in terms of spatialisation quality and the mix does not lead to a homogenous result.

In the particular case where the additional source comes from a spot microphone, high spatial resolution makes its integration into the final recording easier. However, in this case, certain precautions must be taken:

- Additional spot sources must be panned in the same direction as the corresponding phantom images produced by the microphone. Should this not be the case, spatial punctuality is reduced due to the spatial distribution of the phantom image. However, this precaution is no longer necessary when the sound level of the additional sources is substantially louder than that of the source (> 15dB).

- Delays between spot microphones and the multichannel microphone must be adjusted in order to compensate the distance separating the two and the processing delay introduced by the multichannel microphone. An error in delay selection will lead to time differences which will worsen tones upon mixing (comb filter effects).

On the other hand, there are no special recommendations concerning the arrangement of levels between spot microphones and the multichannel microphone. The two most extreme uses are:

- Dominant multichannel microphone with nearby microphones providing a slight presence only.
- Dominant nearby microphones with the multichannel microphone providing a discrete atmospheric complement.

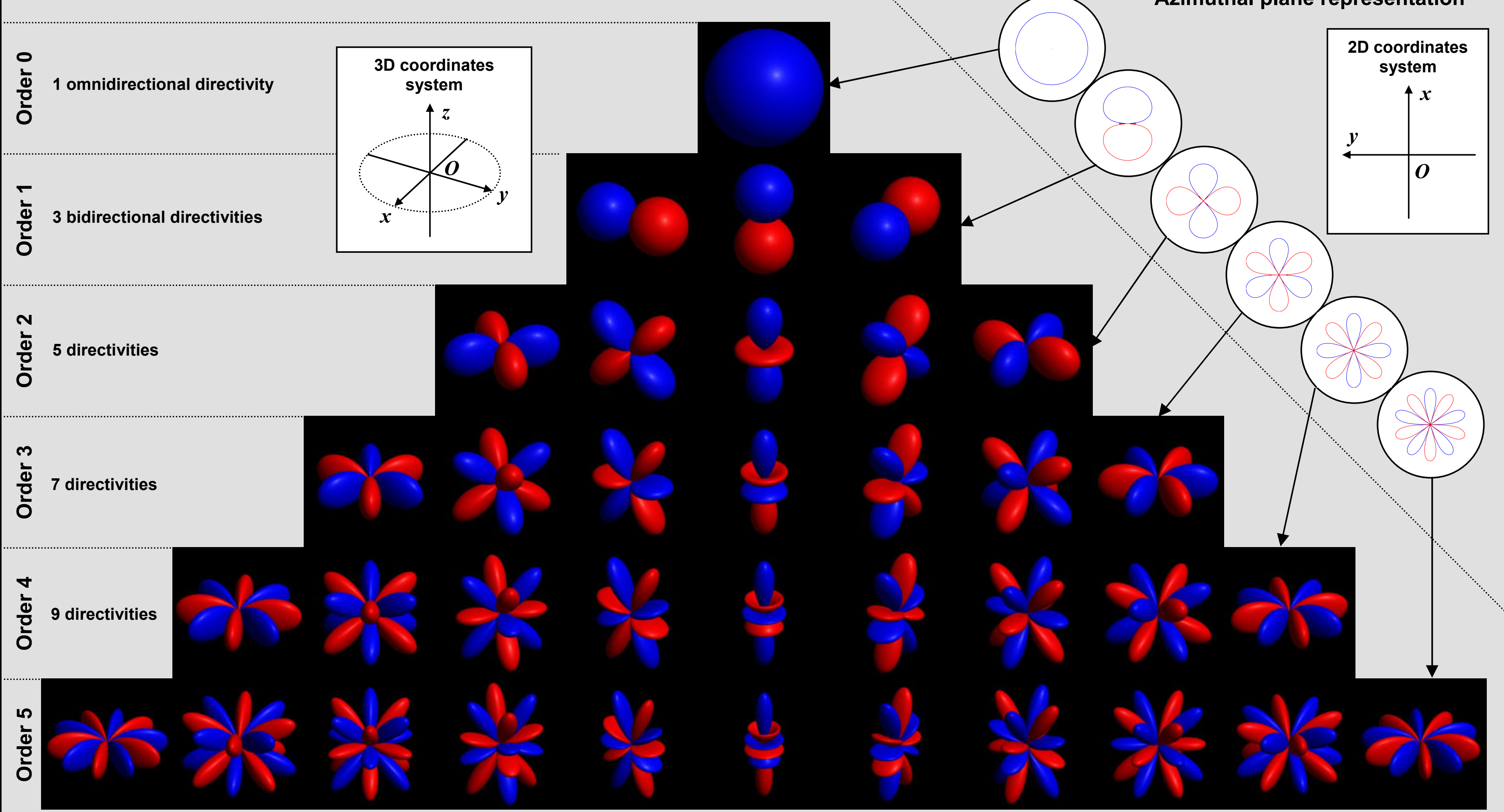
Any arrangement coming between these two extremes is possible. This mainly allows for the relation between the direct field and the reverberated field to be closely controlled in mixing while maintaining the original recorded acoustics.

2.8 Controlling spatialisation outside of the azimuthal zone.

Until now, only the sources placed in the azimuthal zone have been considered. And yet, in an actual sound environment, sound sources are placed in all directions of the sphere (and are not limited to directions in the azimuthal circle). To give an example, a sound source which moves above a microphone must be perfectly recorded in all movements. In the same way, a room will be subject to infinite echoes against walls, producing source images distributed in all directions around the microphone. **Consequently, a multichannel microphone must be optimised for all directions of the sphere.**

High spatial resolution allows for the close control of behaviour outside the azimuthal zone. **Spatial frequencies allow for the equal discrimination of sounds in space, independently of source orientation.** The 36 spatial frequencies of up to order 5 are shown in figure 4-n.

Figure 4-n: Representation of all spatial frequencies of up to order 5.



With these figures, positive spatial frequency lobes are in blue and negative lobes in red.

In an actual sound environment, sounds come from all directions. And yet, in 5.0 configuration, loudspeakers are located in the azimuthal zone. Reproduced sources may only therefore be located in the azimuthal zone. Faced with the limitation to 5.0 format, it is necessary to determine the ideal behaviour for the microphone outside the azimuthal zone. Generally speaking, the ideal arrangement is a source located outside the azimuthal zone when the recording is reproduced in the same azimuthal direction without elevation. Obviously this is overgeneralisation. A source which moves above the microphone from front to back would be reproduced only at the front then it would suddenly switch to the back when going beyond the microphone. This behaviour is not realistic in terms of the actual movement of the source.

A more realistic approach would be:

- Any source in the azimuthal zone is recorded with ideal directivity figures.
- Any source located exactly above (or below) the multichannel microphone is reproduced by the 5 loudspeakers in such a way as to ensure that no orientation is dominant. Such capturing uses no separation and corresponds to sound recording using 5 coinciding omnidirectional microphones.
- Any source located at an intermediate height between the two extreme cases will use an intermediate solution. In other words, recording would gradually become omnidirectional as the source rises.

This recording produces the desired effect when the source moves over the microphone: the reproduced phantom image is brought into the zone, but moves gradually from front to back. Therefore, the concept of ideal directivities, initially defined in the azimuthal zone only, may be extended to all directions surrounding the microphone. Figure 4-o shows **the ideal directivities for optimum multichannel sound recording for all directions in space.**

The directivities required to obtain high spatial resolution 5.0 sound recording present the following characteristics:

- <u>Highly selective</u> for directions close to the azimuthal zone, i.e. for elevations of +/-45° from the azimuthal zone. Again, this spatial selectivity is at the origin of high spatial resolution.

- <u>Asymmetry</u>. The absence of symmetry as compared with the directions of the loudspeaker had already been shown in the azimuthal zone. The absence of axial symmetry appears in an even more obvious manner when we compare behaviour in the azimuthal zone and in elevation from directivities.

Obviously, such ideal directivity respects the essential properties of high spatial resolution for all directions in the zone:

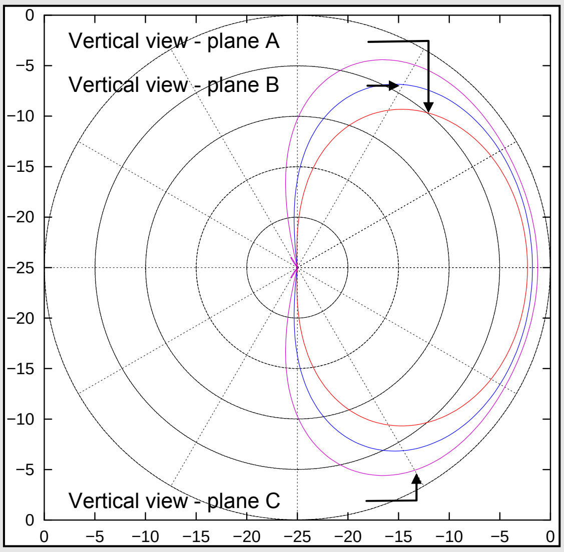
- Optimum channel separation. An exaggerated separation of channels for elevated sources produces unrealistic source movements.
- Increased source punctuality.
- Increasing the size of the sweet spot.
- Maximum respect of tones in all directions
- Presence of phantom images over 360°.
- Easy and good quality down-mix independently of source orientation.
- Easy addition of spot microphones.
- Better control of the coverage angle in the horizontal and vertical planes.

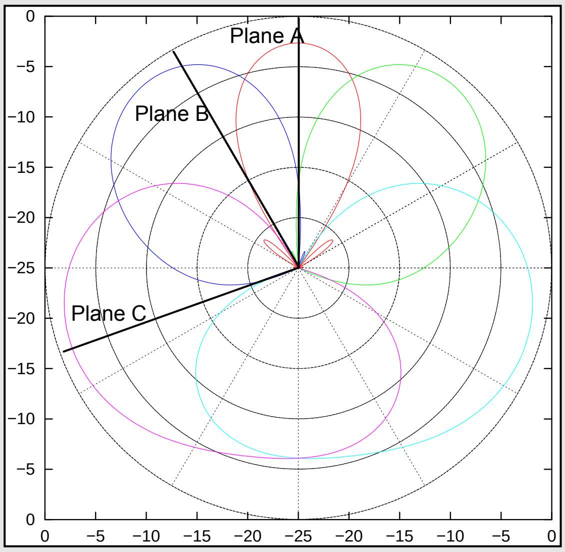
The high spatial resolution sound recording systems developed by **Trinnov** *Audio* are based on 36 spatial frequencies and allow for close control of ideal directivities for all directions in space. **High spatial resolution particularly allows for the use of a totally different approach between the azimuthal zone and elevation.**

Figure 4-o: Ideal directivities for 5.0 sound recording for all directions in space.

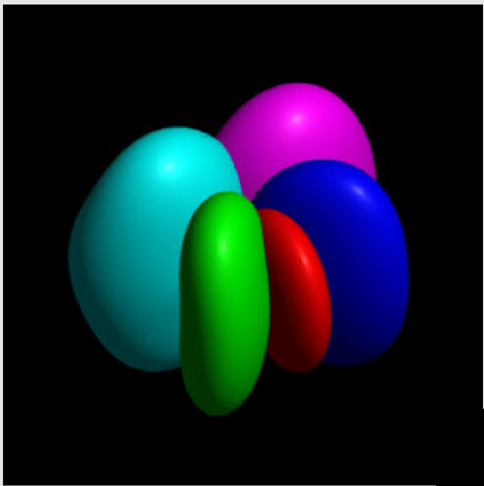
Vertical views according to 5.0 format directions.

Horizontal view

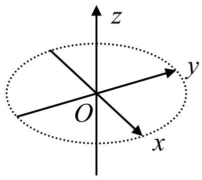




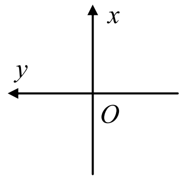
Representation of all directivities for all orientations



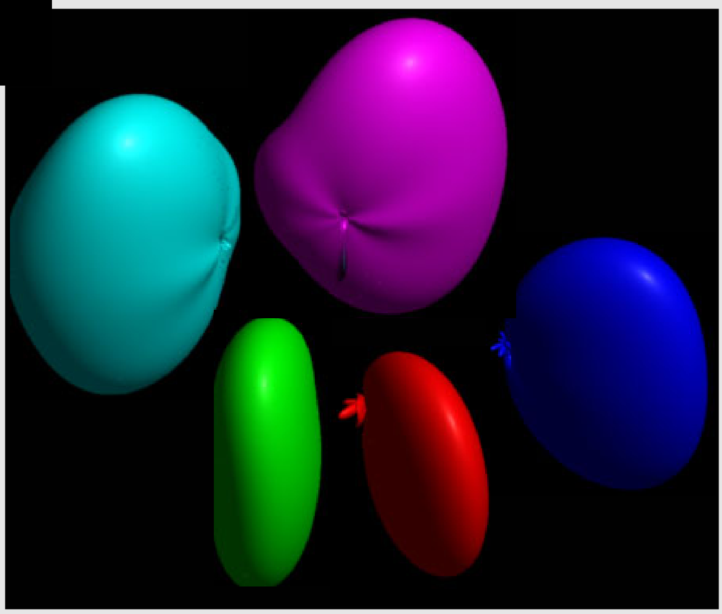
3D coordinates system



2D coordinates system



"Exploded" view of ideal directivities



2.9 Optimum control of the distance factor

When an actual sound source radiates in an actual environment, we usually consider that the acoustic field obtained is a mix between:

- direct field, formed by the first wave which reaches the microphone and possibly by the first reverberations against the walls of the recording venue (mainly those coming from the ground). Sound intensity reduces when the distance between the source and the microphone increases. The sound intensity of the direct field is always oriented in the direction of the source at the position of the microphone.
- reverberated field, made up of the superposition of the infinite reverberations of sound against the walls of the recording venue. The sound intensity of the reverberated field is strictly identical to that of the recording venue. For each of these points, the sound intensity is equally distributed between all directions.

The intensity of the direct field as compared with that of the reverberated field varies depending on the distance between the source and the microphone. The distance for which the direct field and the reverberated field have the same intensity is called critical distance. In general, an analytic sound recording in which the sounds are broken down and sound sources are precisely identified is obtained by placing the microphone at a distance shorter than critical distance.

Unfortunately, often recording conditions set the critical distance at just a few metres. With monophonic sound recording, the intensity of the direct field as compared with the reverberated field may be modified using a directive microphone. To give an example, a cardioid microphone oriented in the direction of the source has a direct field sensitivity 1.73 times higher (4.8dB) than that of the reverberated field. Therefore, the critical distance is artificially increased by factor 1.73 (we call this the distance factor).

Obviously, considerations concerning the critical distance also apply to multichannel sound recording. In this case, the 5 channels must be considered simultaneously in order to determine the overall sensitivity of the system with regard the direct field and the reverberated field. In a traditional recording situation where sources are located in the frontal zone, the frontal channels (L,R,C) reproduce a mix of direct field and reverberated field while the back channels (SL,SR) mainly reproduce reverberated field. An immediate solution for increasing critical distance involves reducing the level of back channels. Unfortunately this solution damages spatial fidelity as lateral images gather on the front right and front left loudspeakers.

Trinnov *Audio* technology allows the overall sensitivity of the multichannel microphone to be controlled depending on the direction of sources, without altering tones or distorting angles. Control is obtained by maintaining all properties of high spatial resolution. This control is possible thanks to the 36 spatial frequencies of up to order 5.

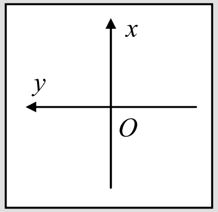
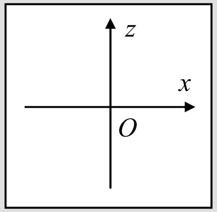
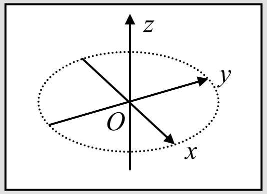
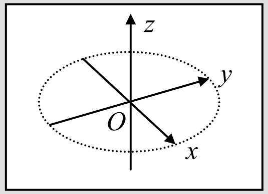
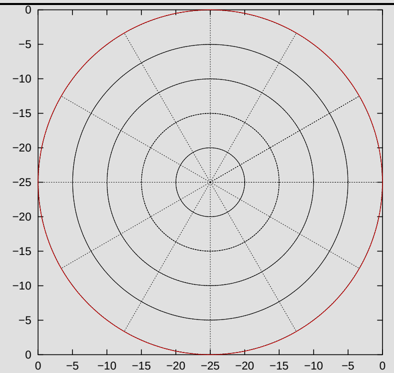
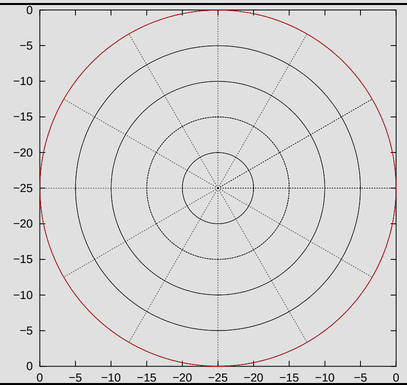
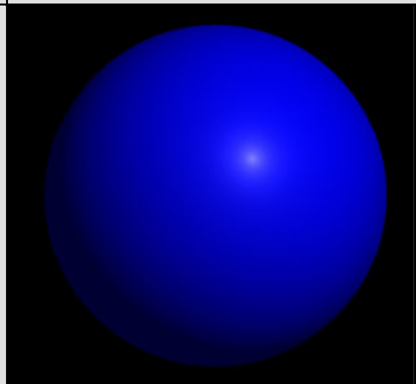
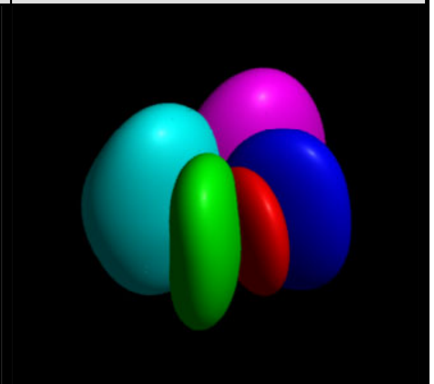
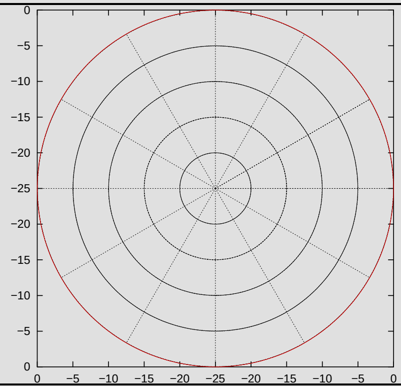
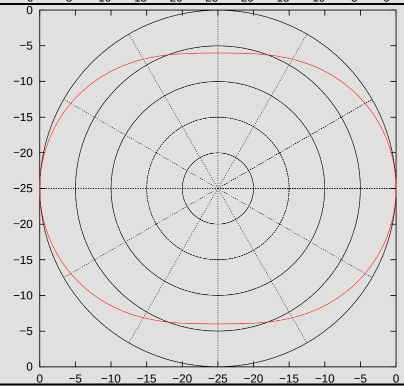
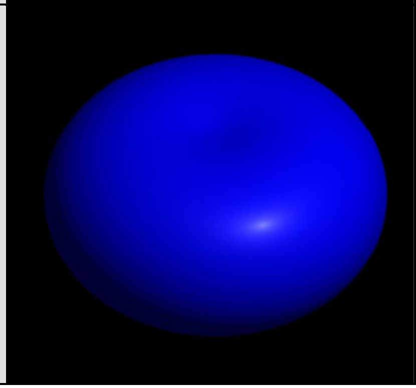
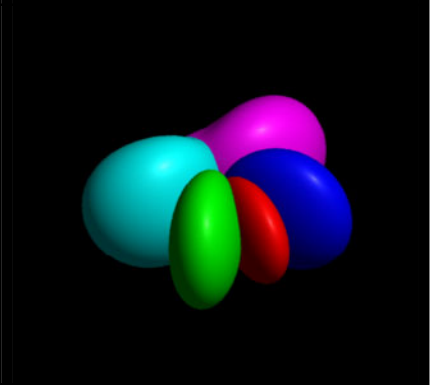
The principle consists of reducing the overall sensitivity of the microphone in directions with no sources and reverberated field only. This reduction in sensitivity affects all ideal directivities in an identical manner. There is therefore no change in the position of phantom images, but sound levels are modified. The intensity of the direct field is maintained while that of the reverberated field is reduced: the distance factor is higher than 1. Several strategies may be envisaged depending on the distribution of sources surrounding the microphone:

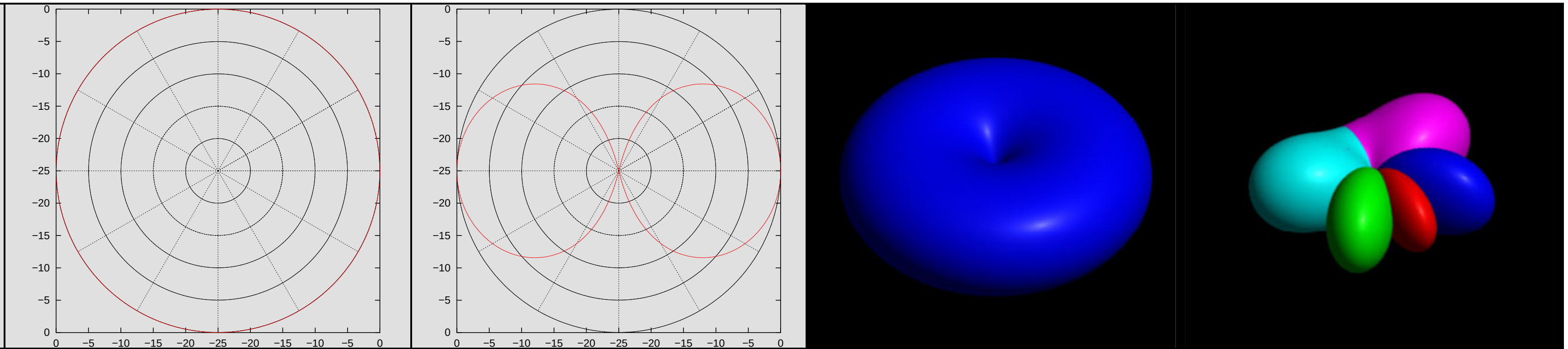
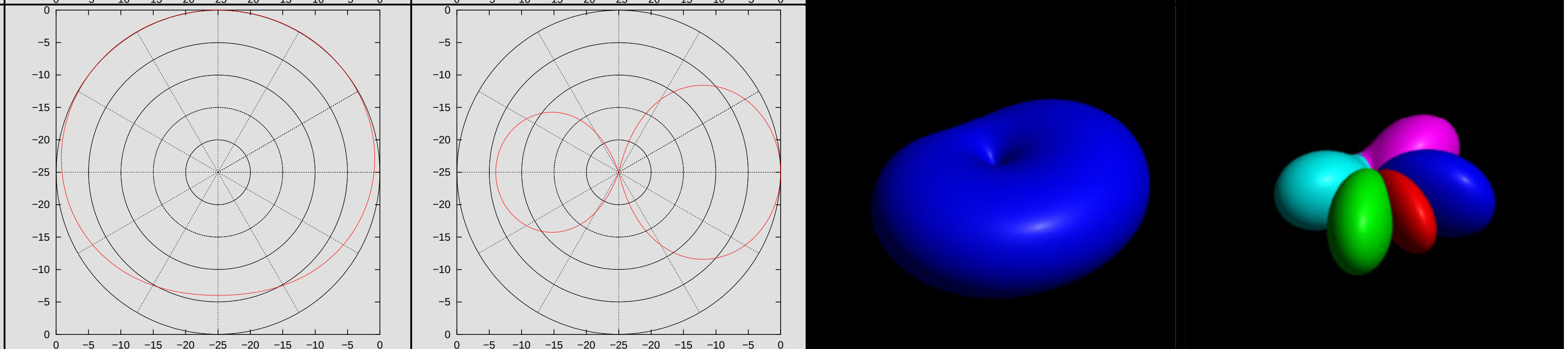
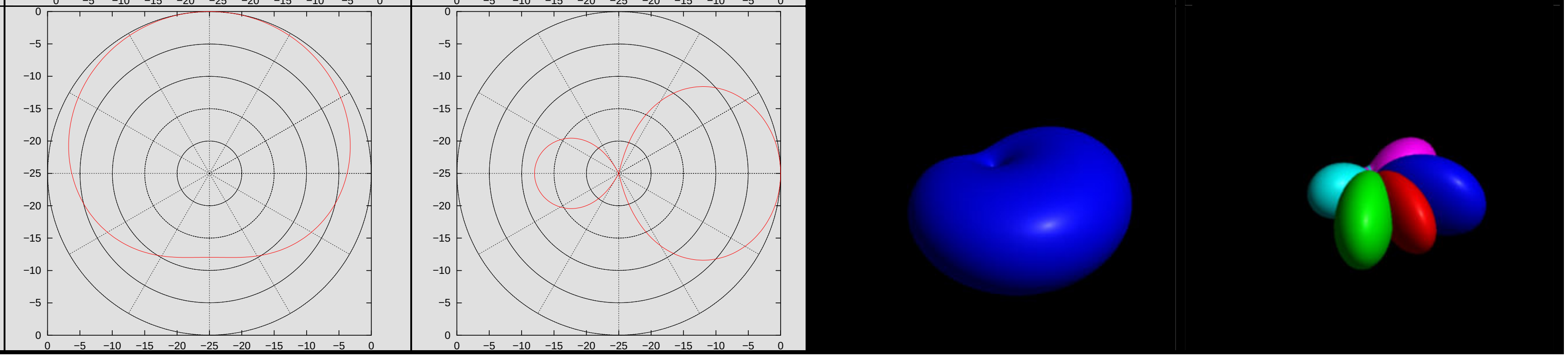
- Sources surrounding the microphone over 360° of the azimuthal zone and no substantial elevation (+/- 40° in relation to the azimuthal zone). In this case, the distance factor may be increased by reducing the overall sensitivity of the microphone for directions with high elevation (above or below the microphone).

- Sources located in the frontal zone (+/-90°) and no substantial elevation (+/- 40° in relation to the azimuthal zone). In this case, the distance factor may be increased by reducing the sensitivity for sources in elevation (above or below the microphone) and located in the back zone for the microphone.

Figure 4-p shows the control of critical distance provided by high spatial resolution.

Figure 4-p: Control of critical distance.

	Horizontal view	Vertical view	Overall sensitivity for all directions	Ideal directivities for all directions
				
Full Omni Directivity factor: 0 dB Distance factor: 1				
Oblate Omni Directivity factor: 0 dB Distance factor:				

Omni without ceiling Directivity factor: 2.7 dB Distance factor: 1.37	
Back reduction Directivity factor: 4.2 dB Distance factor: 1.63	
Back reduction Directivity factor: 5.9 dB Distance factor: 1.98	

As previously highlighted, the 5 channels of a 5.0 high spatial resolution sound recording are aligned and combine perfectly in the centre of the restitution system. This operation is the equivalent of a mono down-mix. Should the critical distance not be artificially modified, the mono down-mix will correspond exactly to a mono omnidirectional sound recording. In other cases, the mono down-mix will correspond to sound recording undertaken by a microphone whose directivity corresponds to the overall sensitivity of the multichannel microphone. Such directivity requires an extended spatial frequency band and has no axial symmetry. For reasons similar to those put forward in section 1.4, this directivity cannot be obtained with one single microphone. Only **Trinnov** *Audio* high spatial resolution technology allows directivities to be controlled in this way. With a directivity factor of 5.9 dB, performances may be higher than those obtained with a cardioid (4.8 dB) while the sounds from the back zone remain audible (reduction of 12 dB).

2.10 From the coverage angle to angle distortion

Up to now, we have described multichannel microphones with perfect spatial fidelity. Reproduced phantom sources are exactly located in the direction of the recorded sources. And yet, it is sometimes desirable to use a different distribution. To give an example with the case of sound recording close to a wide sound zone, it may be preferable to squeeze the frontal zone.

With stereophony, the concept of coverage angle was developed in order to show the angular opening for which recorded sources are located between two loudspeakers. However, the coverage angle provides no information on the manner in which sources are distributed in the stereophonic space. The concept of angular distortion is based on this idea, but goes on to complete it. The angular distortion curve specifies the direction of the phantom image for each corresponding direction in space.

The idea of angular distortion is based on order 1 spatial frequencies. According to the meta-theory of aural perception by Michael Gerzon:

- the order 0 spatial frequency (omnidirectional component) of the acoustic field is representative of source tone.
- the 3 order 1 spatial frequencies (bidirectional components) of the acoustic field are representative of the spatial position of the source. However, these frequencies provide no information on the other characteristics of sources such as size, form, orientation, …

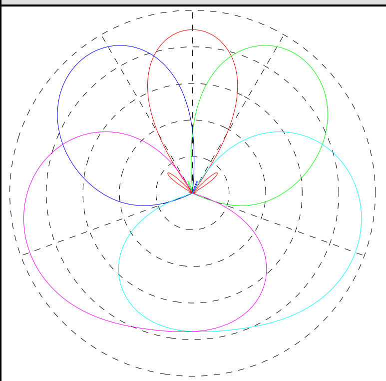
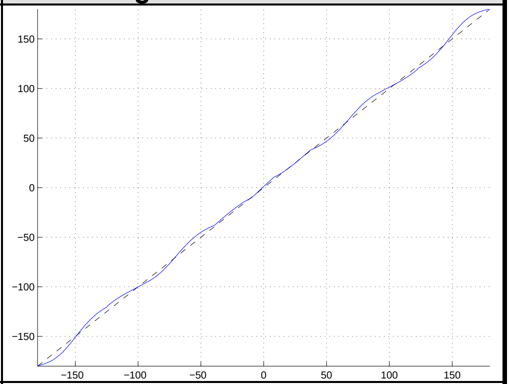
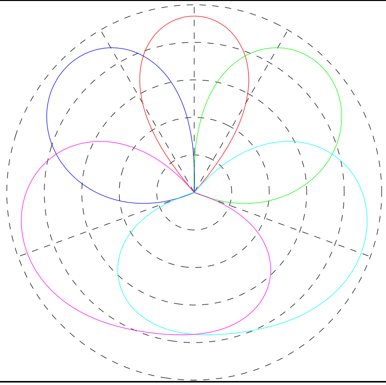
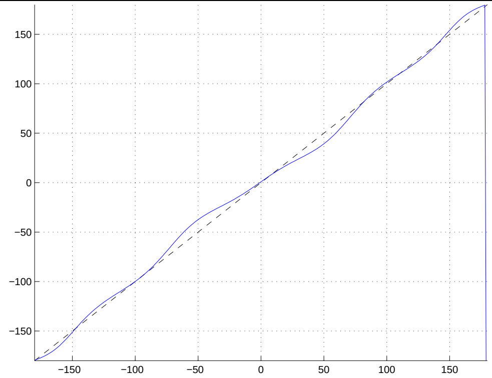
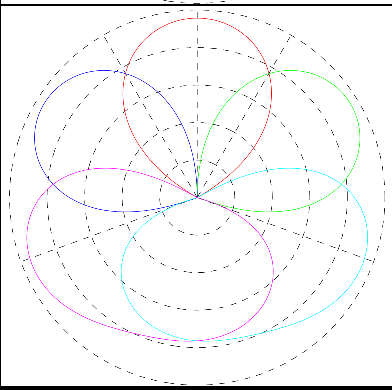
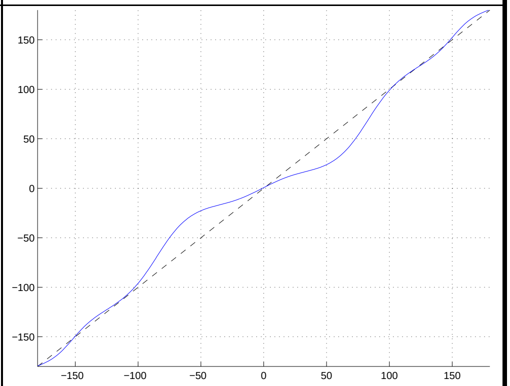
As previously mentioned in section 2.4., the comparison of omnidirectional components between the recorded field and the reproduced field allows for the determination of tone alteration added by the multichannel microphone. In the same way, comparison of bidirectional components between the recorded field and the reproduced field allows for the determination of changes in position added to sound sources. In other words, **the analysis of order 1 spatial frequencies allows angular distortion to be carefully defined**.

On the contrary, the technology developed by **Trinnov** *Audio* allows spatial frequencies to be controlled, particularly order 1 spatial frequencies. In this way, this technology enables the careful control of angular distortion of high spatial resolution multichannel microphones. Control of angular distortion is shown in figure 4-q, for angles covering 30° (no distortion), 40° (slight distortion) and 60° (substantial squeezing of the frontal zone).

2.11 Better robustness in processing

This is an immediate consequence of high channel separation. Channel modification does not affect phantom images other than on each side of the loudspeaker corresponding to the modified channel. The rest of the sound zone is not affected by this modification. In this way, high spatial resolution provides increased robustness for usual processing, such as level correction, equalising, dynamic compression or the coding of active matrixes (pro-logic type).

Figure 4-q: Control of angular distortion.

Sound recording	Ideal directivities	Angular distortion curve
No angular distortion		
Coverage angle of 40°		
Coverage angle of 60°		

3 Scientific basis of High Spatial Resolution: controlling the acoustic field

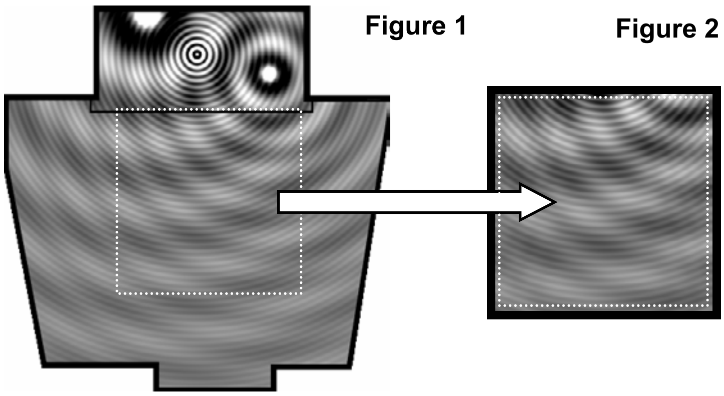
3.1 Scientific basis for the technology

In parallel to traditional stereophony-based techniques, certain research laboratories are looking into new approaches, particularly *IRCAM*, *Université de Delft*, *France Télécom R&D* and **Trinnov** *Audio*. **The central idea is to consider the physical nature of sound i.e. the acoustic field.** The technologies developed are known as *Wave Field Synthesis* (*WFS*) or *High Order Ambisonics* (*HOA*). These technologies are based on fundamental acoustics: Huygens's principle (1690) and the Fourier-Bessel series (1870). This is a sort of absolute approach which allows for the acoustic field in the recording venue to be recorded in a wide area and reproduced in a perfectly identical manner in the restitution venue. In this way, if we are in the sweet spot or in the recording venue, no difference is noticeable, the aural experience is totally identical. Unfortunately, these systems are not used as hundreds of microphones and loudspeakers are required.

Figure 4-r: Control of the acoustic field.

Figure 1: In this concert hall seen from above, three sound sources create an acoustic field in the middle of which the listeners are placed.

Figure 2: In this auditorium, the acoustic field captured during the concert is perfectly reproduced. Listeners are thus placed in the middle of the same acoustic field as the concert.



However, **this scientific basis confirms the concept of high spatial resolution in an extensive and objective manner.** HOA technology is precisely based on the breakdown of the acoustic field into spatial frequencies known as Fourier-Bessel functions (made up of Bessel functions and spherical harmonics). The impact of this technique is identical to that provided by the breakdown of a signal into its frequencies: tones may be entirely transformed, or, on the contrary, perfectly maintained. In the same way, HOA is an effective means of controlling the acoustic field: tones and source positions may be entirely transformed, or, on the contrary, perfectly maintained.

3.2 High spatial resolution recording - Sampling principle

As mentioned in section 1.4, the directivities required for recordings with high spatial resolution are not provided naturally by current microphones. To obtain the required performances, it is necessary to recombine them using complex linear processing: full matrix filtering. A high spatial resolution multichannel microphone operates in two stages:

- Stage 1: acquisition of information on the acoustic field using an array of standard microphones.
- Stage 2: processing, based on full matrix filtering allowing the required directivities to be synthesised.

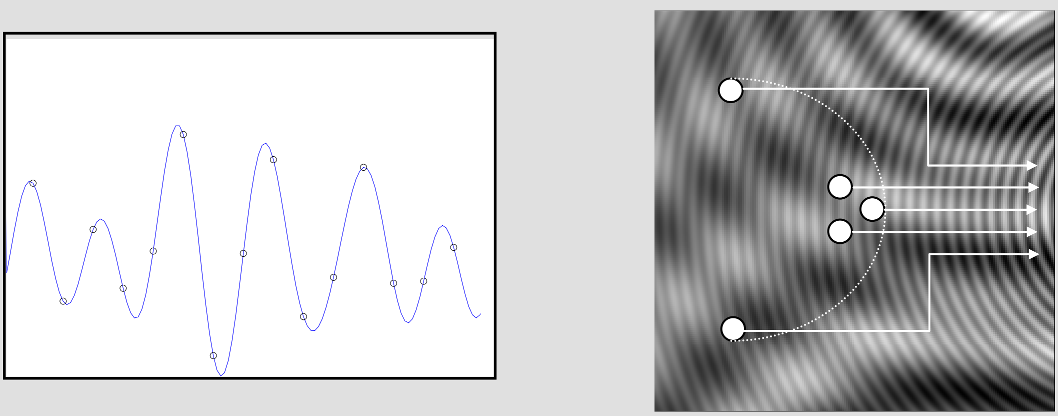
The acquisition stage may be interpreted as a spatial sampling stage of the acoustic field and is directly based on the temporal sampling method used for signals.

A signal (whose frequency band is limited) may only be observed at certain moments without losing information. The initially continuous temporal signal is therefore perfectly represented by a series of values known as samples. This is described as temporal signal sampling.

In the same way, an acoustic field may only be observed in certain positions without losing information. The initially continuous acoustic field is therefore perfectly represented by a series of signals known as samples. This is described as spatial sampling of the acoustic field.

In reality, a microphone array takes precise measures of the acoustic field at different positions and represents a sampler of the acoustic field. All measures undertaken characterise the spatial variations of the acoustic field. This principle is shown in figure 4-s.

Figure 4-s: Spatial sampling of the acoustic field.



3.3 High spatial resolution recording - Encoding principle

Encoding is an operation which involves determining the spatial frequencies of the acoustic field. When the microphone array is exposed to an acoustic field of which all aspects are known, the laws of physics allow us to calculate samples, i.e. the signals emitted by the network microphones.

However, in a recording situation, the situation is exactly the opposite: the microphone array emits signals which are known and the aim is to obtain information on the initial acoustic field. In the particular case of high spatial resolution recording, the aim is to measure the spatial frequencies of the acoustic field up to order 5. **In this way, the encoding operation corresponds to the exact opposite of the spatial sampling process of the acoustic field. A high spatial resolution recording is obtained by reversing the sampling process of the acoustic field using full matrix filtering.**

A detailed explanation of this technique can be found in the publication "*A New Comprehensive Approach of Surround Sound Recording*" presented at the 114^{th} AES Convention (preprint 5717). To obtain order 5 spatial resolution, this system requires a minimum of 36 microphones (usually 48).

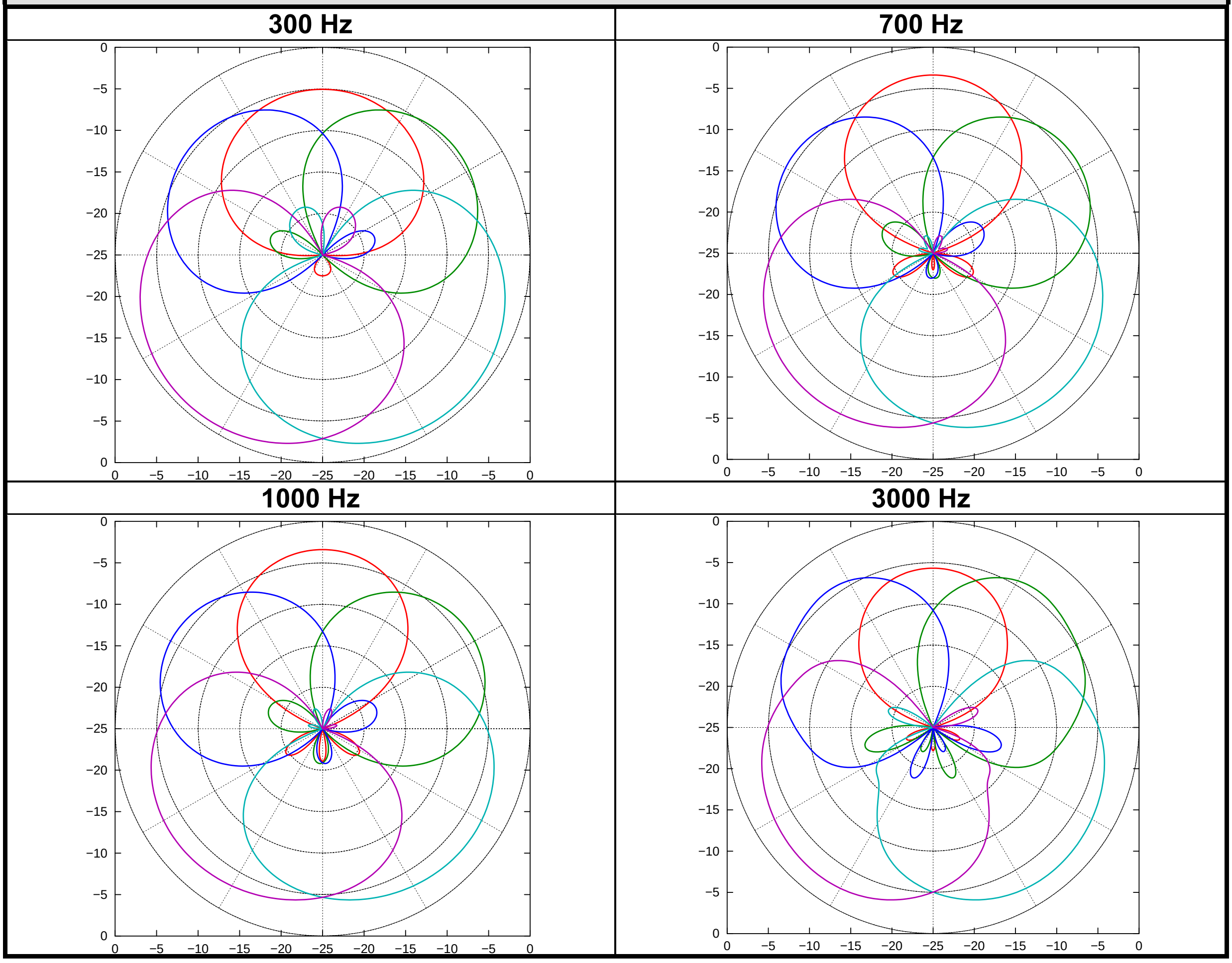
3.4 Capture in ∆T and transmission in ∆I

Thanks to the work of **Trinnov** *Audio*, high spatial resolution may be obtained with an array of **8 omnidirectional microphones and one processing unit**.

With this system, a purely ∆T approach is used to capture the acoustic field. **The information is taken uniquely thanks to the distance between captors**, no use is made of directional characteristics. The microphone array layout is optimised to obtain a maximum of information on the acoustic field as effectively as possible. **Trinnov** *Audio*'s works have proved that, in view of the low spatial selectivity of current microphones, a non-coinciding network captures better than a coinciding network (*114th AES Convention, preprint 5717*). In other terms, the sampling of an acoustic field by a non-coinciding microphone array is more effective than sampling with a coinciding network.

Thanks to acoustic field control-based processing, the processing unit determines the 5 channels in 5.0 format leading to optimum reproduction of the acoustic field. Processing is based on 40 filters at 1024 frequencies and produces optimum reconstitution of each channel using 8 omnidirectional microphones. The entire system (network + processing unit) acts as 5 coinciding microphones providing high spatial resolution sound recording. Typical performances of the **Trinnov** *Audio* HSR 5.0 microphone are shown in figure 4-t. Information is coded in sound intensity variation (ΔI) only in the multichannel signal. This type of coding is particularly suitable for the transmission and restitution of multichannel sound.

Figure 4-t: Typical performances of the Trinnov *Audio* HSR 5.0 microphone



3.5 Trinnov SRP: a platform for high spatial resolution.

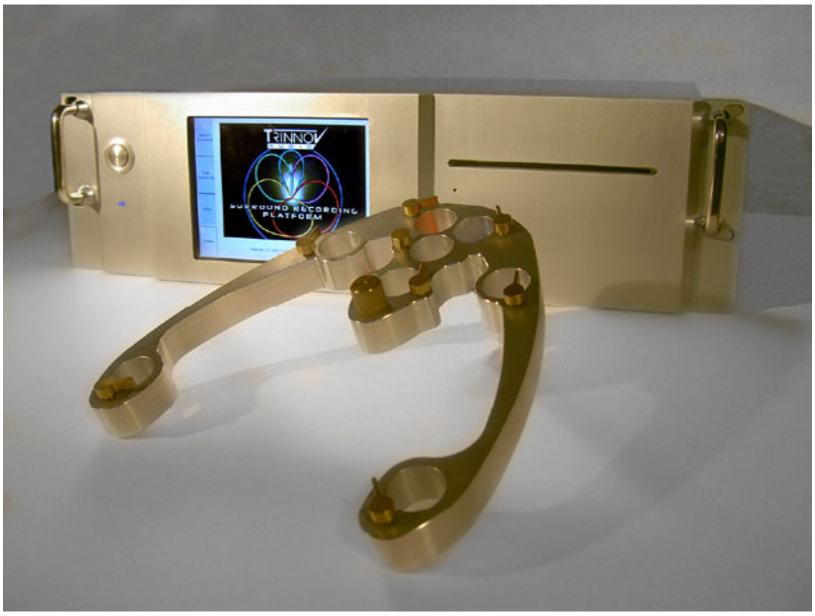
High spatial resolution technology provides new prospects in terms of multichannel sound recording. By using an original structure made up of a microphone array and specific digital processing, this technology is the starting point for a new generation of multichannel microphones.

Trinnov *Audio* proposes a flexible and adaptable solution*: Trinnov SRP* (*Surround Recording Platform*). More than a simple multichannel microphone, this is an entire adaptable platform for high spatial resolution sound recording. The *SRP* is currently the first high spatial resolution 5.0 multichannel microphone. This platform is designed to accept future high spatial resolution microphones to be proposed in the form of *HSR* (*High Spatial Resolution*) packages. A package is made up of a CD for updating SRP software and possibly a new microphone array.

The first available package is the *HSR 5.0/Omni-8* including:

- the *Omni-8* array accepting 8 omnidirectional studio microphones. Thanks to a set of rings, you can adapt your favourite studio microphones.
- the *HSR 5.0* full matrixing filter bank required for use of the Omni-8 array.
- monitoring tools for high spatial resolution sound recording.
- distance factor management tools.

Trinnov SRP equipped with *HSR 5.0/Omni-8*



Trinnov *Audio*
30-32 avenue de la République
94815 Villejuif Cedex
France

Tel : +33 1 46 81 27 56
Fax :+33 1 46 81 77 27
Mob : +33 6 16 81 62 87
www.trinnov.com – contact@trinnov.com

For more information on Trinnov SRP, please contact:

Sebastien Montoya, Product Manager
sebastien.montoya@trinnov.com
Tel : +33 6 74 09 95 38

Arnaud Laborie, General Manager
arnaud.laborie@trinnov.com
Tel: +33 6 16 81 62 87
Fax :+33 1 46 81 77 27