Ambisonics Explained: A Guide for Sound Engineers

Ambisonics解释：声音工程师指南

October 10, 2017

<https://www.waves.com/ambisonics-explained-guide-for-sound-engineers>

Need to mix audio for 360 videos or VR projects and not sure where to start? Get a jumpstart with this basic guide to Ambisonics B-format, the most widely used audio format for 360 applications.

需要为360视频或VR项目混合音频，但不确定从哪里开始？ 快速入门此Ambisonics B格式基本指南，这是360应用程序中使用最广泛的音频格式。



In recent years, VR game engines, 3D installations, and 360° videos on social media such as YouTube and Facebook have all become dependent on full-immersion, three-dimensional, 360-degree audio. If you’re a sound engineer, you are more likely than ever to be working on projects that involve converting, mixing, panning and playing back 360 audio.  
  
The most popular audio standard for handling and delivering such audio is called Ambisonics, and professional [Ambisonics audio tools](https://www.waves.com/hardware/360-ambisonics-tools) adapted to the familiar workflow of audio engineers are now available. The information on this page will help you understand the basic concepts of Ambisonics and the basic workflow of handling 360-degree audio projects in Ambisonics format.

近年来，VR游戏引擎，3D装置以及YouTube和Facebook等社交媒体上的360°视频已经完全依赖于全浸入式，三维360度音频。 如果您是音响工程师，那么您比以往任何时候都更有可能从事涉及转换，混合，平移和播放360音频的项目。

用于处理和传递此类音频的最流行的音频标准称为Ambisonics，现在可以使用适合音频工程师熟悉的工作流程的专业Ambisonics音频工具。 本页上的信息将帮助您了解Ambisonics的基本概念以及处理Ambisonics格式的360度音频项目的基本工作流程。

# What is Ambisonics? 什么是Ambisonics

Ambisonics is a method for recording, mixing and playing back three-dimensional 360-degree audio. It was invented in the 1970s but was never commercially adopted until recently with the development of the VR industry which requires 360° audio solutions.  
  
The basic approach of Ambisonics is to treat an audio scene as a full 360-degree sphere of sound coming from different directions around a center point. The center point is where the microphone is placed while recording, or where the listener’s ‘sweet spot’ is located while playing back.  
  
The most popular Ambisonics format today, widely used in VR and 360 video, is a 4-channel format called **Ambisonics B-format**, which uses as few as four channels (more on which below) to reproduce a complete sphere of sound.

Ambisonics是一种用于录制，混合和播放3D 360度三维音频的方法。 它是在1970年代发明的，但是直到最近需要360°音频解决方案的VR行业的发展才被商业应用。

Ambisonics的基本方法是将音频场景视为来自中心点不同方向的完整360度声音范围。 中心点是录音时放置麦克风的位置，或播放时聆听者的“最佳位置”。

如今，在VR和360视频中广泛使用的最受欢迎的Ambisonics格式是一种称为Ambisonics B格式的4通道格式，该格式使用少至四个通道（下面有更多通道）来再现完整的声音范围。

## Why Ambisonics, why now? Ambisonics vs. Surround

Ambisonics 360 audio is sometimes confused with traditional surround sound technologies. But they are not the same, and there are major differences between them. And there are reasons why Ambisonics, rather than classic surround formats, has been adopted as the technology of choice for emerging VR and 360 applications.  
  
Traditional surround technologies are more immersive than simple two-channel stereo, but the principle behind them is the same: they all create an audio image by sending audio to a **specific, pre-determined array of speakers**. Stereo sends audio to two speakers; 5.1 surround to six; 7.1 to eight; and so on.  
  
By contrast, Ambisonics does not send audio signal to any particular number of speakers; it is “speaker-agnostic.” Instead, **Ambisonics can be decoded to *any* speaker array** (more on which below). Ambisonic audio represents a full, uninterrupted sphere of sound, without being restricted by the limitations of any specific playback system.  
Ambisonics 360音频有时会与传统的环绕声技术混淆。 但是它们并不相同，它们之间存在主要差异。 并且为什么将Ambisonics而不是经典的环绕声格式用作新兴VR和360应用程序的首选技术是有原因的。

传统的环绕声技术比简单的两声道立体声技术更具沉浸感，但是其背后的原理是相同的：它们都通过将音频发送到特定的预定扬声器阵列来创建音频图像。 立体声将音频发送到两个扬声器, 5.1环绕至6通道, 7.1至8通道等等。

相比之下，Ambisonics不会将音频信号发送到任何特定数量的扬声器。 这是“与扬声器无关的”。 取而代之的是，可以将Ambisonics解码为任何扬声器阵列（下面将详细介绍）。 Ambisonic音频代表完整，连续的声音范围，不受任何特定播放系统的限制。

All this helps explain why Ambisonics has become standard in 360 video and VR:

* Traditional surround formats can provide good imaging when static; but as the sound field rotates, the sound tends to ‘jump’ from one speaker to another. By contrast, Ambisonics can create a smooth, stable and continuous sphere of sound, even when the audio scene rotates (as, for example, when a gamer wearing a VR headset moves her head around). This is because Ambisonics is not pre-limited to any particular speaker array,
* Traditional surround speaker systems are usually ‘front-biased’: information from the side or rear speakers is not as focused as the sound from the front. By contrast, Ambisonics is designed to spread the sound evenly throughout the three-dimensional sphere.
* Finally, whereas traditional surround systems have various difficulties representing sound beyond the horizontal dimension, Ambisonics is designed to deliver a full sphere complete with *elevation*, where sounds are easily represented as coming from above and below as well as in front or behind the user.

所有这些都有助于解释为什么Ambisonics成为360视频和VR的标准：

* 传统的环绕声格式在静态时可以提供良好的成像效果；但是随着声场的旋转，声音往往会从一个扬声器“跳”到另一个扬声器。相比之下，Ambisonics可以创建平滑，稳定和连续的声音范围，即使在音频场景旋转时（例如，当戴着VR头戴式耳机的游戏者摇头时）。这是因为Ambisonics并不限于任何特定的扬声器阵列，
* 传统的环绕扬声器系统通常是“前偏”的：侧扬声器或后扬声器的信息不如前扬声器的声音集中。相比之下，Ambisonics旨在将声音均匀地分布在三维球体上。
* 最后，尽管传统的环绕声系统在表示声音时超出了水平方向，但存在各种困难，而Ambisonics旨在提供完整的球体和高程，其中声音很容易表示为来自使用者的上方和下方以及前方或后方。

## 4 channels = Full 360 degrees

At this point you may be wondering what all this means for a working audio engineer: how exactly is Ambisonics recorded, mixed and played back, if it does not eventually route to individual speakers? But before we jump ahead to the practical aspects, let’s spend a little more time on a more basic theoretical question:  
  
How does Ambisonics represent an entire 360-degree soundfield with as few as four channels?  
  
Let's take a look at the most basic (and today the most widely used) Ambisonics format, the **4-channel B-format**, also known as **first-order** Ambisonics B-format.  
  
The four channels in first-order B-format are called **W, X, Y and Z**. One simplified and not entirely accurate way to describe these four channels is to say that each represents a different directionality in the 360-degree sphere: center, left-right, front-back, and up-down.

在这一点上，您可能想知道这对一个工作的音频工程师意味着什么：如果Ambisonics最终没有路由到各个扬声器，那么它们如何被准确地记录，混合和播放？ 但是，在我们进行实际操作之前，让我们花一些时间在一个更基本的理论问题上：

Ambisonics如何代表多达四个声道的整个360度声场？

让我们看一下最基本的（今天也是使用最广泛的）Ambisonics格式，即4通道B格式，也称为一阶Ambisonics B格式。

一阶B格式的四个通道分别称为W，X，Y和Z。描述这四个通道的一种简化且并非完全准确的方式是说每个通道在360度范围内代表不同的方向性：中心， 左右，前后和上下。  
  
A more accurate explanation is that each of these four channels represents, in mathematical language, a different **spherical harmonic component** – or, in language more familiar to audio engineers, a **different microphone polar pattern pointing in a specific direction**, with the four being coincident (that is, conjoined at the center point of the sphere):

* **W** is an omni-directional polar pattern, containing all sounds in the sphere, coming from all directions at equal gain and phase.
* **X** is a figure-8 bi-directional polar pattern pointing forward.
* **Y** is a figure-8 bi-directional polar pattern pointing to the left.
* **Z** is a figure-8 bi-directional polar pattern pointing up.

更为准确的解释是，这四个通道中的每个通道都用数学语言表示不同的球谐分量–或用音频工程师更熟悉的语言表示指向特定方向的不同麦克风极性模式，这四个通道是重合的（ 也就是说，在球体的中心点相连）：

* W是一种全向极性模式，包含来自所有方向的具有相同增益和相位的球形声音。
* X是向前指向的8字形双向极性图形。
* Y是指向左侧的8字形双向极性图形。
* Z是向上的8字形双向极性图形。

An even fuller explanation will sound familiar to anyone acquainted with omni and figure-8 bi-direction microphones. Take the X channel described above. Like any figure-8 microphone, it has a positive side and a negative (inverse phase) side. While the X channel’s figure-8 polar pattern points forwards, its negative side points backwards. The resulting audio signal on the X channel contains all the sound that is in the front of the sphere with positive phase, and all the sounds from the back of the sphere with negative phase. Also, as with figure-8 microphones, the gain picked up for each direction is different: the signal directly in front or behind will be picked up with full gain, but as you move away from this bi-directional axis the gain drops, until at exactly 90 degrees to the figure-8 you will get zero gain.

熟悉全向和Figure-8双向麦克风的任何人都将听到更充分的解释。 采取上述X频道。 像任何8字形麦克风一样，它具有正极性和负极性（反相）。 X通道的8字形极性图形指向前方，而其负侧指向后方。 X通道上的最终音频信号包含处于正相的球体前部的所有声音，以及来自负相的球体后部的所有声音。 另外，与图8麦克风一样，每个方向上采集的增益也不同：将以全增益采集前或后的信号，但是当您远离该双向轴时，增益下降，直到 与图8完全成90度时，您将获得零增益。  
  
The same goes for the Y and Z channels: The Y channels pick up the left side of the sphere with positive phase and the right side with negative phase. The Z channel picks up the top side of the sphere with positive phase and the bottom with negative phase. This way, by means of differential gain and phase relations, the four channels combined represent the entire three-dimensional, 360-degree sphere of sound.  
Y和Z通道也是如此：Y通道以正相位拾取球体的左侧，而以负相位拾取球体的右侧。 Z通道以正相位拾取球体的顶部，以负相位拾取球体的底部。 这样，借助差分增益和相位关系，组合的四个通道代表了整个三维360度的声音范围。

All this, by the way, should sound very familiar to any audio engineer with an understanding of mid-side (M/S) stereo processing. Recording M/S requires two coincident microphones:

* **M**: an omni mic for the mid (analogous to the W channel in B-format).
* **S**: a figure-8 mic for the sides (analogous to the Y channel in B-format).

Together, the M and S channels capture the entire L/R stereo field via differences in gain and phase on the side channel (where L=M+S, R=M-S). The exact same principle holds for the WXYZ channels of B-format, only with two extra channels providing depth (X) and elevation (Z).

顺便说一下，所有这些对于熟悉中端（M / S）立体声处理的音频工程师来说都应该非常熟悉。 录制M / S需要两个重合的麦克风：

•M：中音（类似于B格式的W通道）。

•S：侧面的8字形麦克风（类似于B格式的Y通道）。

M和S通道一起通过侧通道上的增益和相位差（其中L = M + S，R = M-S）捕获整个L / R立体声场。 B格式的WXYZ通道具有完全相同的原理，只有两个额外的通道提供深度（X）和高程（Z）。

## AmbiX vs. FuMa

It is also worth mentioning at this point that there are two conventions within the Ambisonics B-format standard: **AmbiX** and **FuMa**. They are quite similar, but not interchangeable: they differ by the sequence in which the four channels are arranged, with AmbiX, for example, arranged WYZX instead of WXYZ. The [Waves Ambisonics plugins](https://www.waves.com/hardware/360-ambisonics-tools) use the AmbiX format – which is why on the plugin interfaces you’ll see the channels ordered WYZX. To enable you to move back and forth between AmbiX and FuMa, the Waves B360 Ambisonics Encoder plugin includes AmbiX-to-FuMa and FuMa-to-AmbiX convertors.)

在这一点上还值得一提的是，Ambisonics B格式标准中有两个约定：AmbiX和FuMa。 它们非常相似，但不能互换：它们的区别在于四个通道的排列顺序，例如，使用AmbiX排列的WYZX而不是WXYZ。 Waves Ambisonics插件使用AmbiX格式-这就是为什么在插件界面上会看到频道顺序为WYZX的原因。 为了使您能够在AmbiX和FuMa之间来回移动，Waves B360 Ambisonics编码器插件包括AmbiX到Fuma和FuMa到AmbiX转换器。）



Waves B360 Ambisonics Encoder: WYZX channels on the right-hand side

## First-order to sixth-order Ambisonics 一到六阶

Before we get back to the more practical aspects of working with Ambisonics, it is also worth noting that the 4-channel format described above is only a simple, **first-order** form of B-format, which is what most Ambisonics microphones and playback platform support today. While even first-order B-format provides higher-resolution spatial immersion than traditional surround technologies, higher-order B-format audio can provide even higher spatial resolutions, with more channels providing more different polar patterns. Thus, second-order Ambisonics uses 9 channels, third-order Ambisonics jumps up to 16 channels, all the way up to sixth-order Ambisonics with 49 channels.

在回到使用Ambisonics的更实际的方面之前，还应注意，上述4通道格式只是B格式的一种简单的一阶形式，这是大多数Ambisonics麦克风和播放平台所采用的格式 今天支持。 尽管即使是一阶B格式也能提供比传统环绕声技术更高的分辨率空间沉浸感，但高阶B格式音频也可以提供更高的空间分辨率，更多的通道可以提供更多不同的极性模式。 因此，二阶Ambisonics使用9个通道，三阶Ambisonics跳至16个通道，一直到具有49个通道的六阶Ambisonics。

# Recording, encoding and playing back Ambisonics B-format 记录，编码和播放Ambisonics B格式

Let’s get back now to the more practical aspects of working with first-order B-format. An audio engineer delivering Ambisonics audio for a VR or 360 project could be dealing with one of two basic scenarios:

* The entire session may already be in Ambisonics B-format (for example, it may have been originally recorded using an Ambisonics microphone); or
* The session may be in traditional surround, in which case you would need to convert it to Ambisonics; or perhaps it includes several separate mono or stereo elements from which you want to create a new Ambisonics mix, in which case, again, you would need to convert the tracks and also position them in the final 360° mix.

 现在让我们回到使用一阶B格式的更实际的方面。 为VR或360项目提供Ambisonics音频的音频工程师可能会处理以下两种基本情况之一：

* 整个会话可能已经是Ambisonics B格式的（例如，它最初可能是使用Ambisonics麦克风录制的）；
* 要么会议可能在传统环境中，在这种情况下，您需要将其转换为Ambisonics。 或者它可能包含几个单独的单声道或立体声元素，您要从中创建新的Ambisonics混音，在这种情况下，您同样需要转换音轨，并将它们放置在最终的360°混音中。

## Recording Ambisonics 录制

An Ambisonics recording microphone is built of four microphone capsules encased closely together. These capsules are cardioid polar patterns, and the signals they record are usually referred to as “Ambisonics A-format.” The A-format is then transformed to B-format by a simple matrix to the WXYZ channels.

Ambisonics录音麦克风由四个紧密封装在一起的麦克风胶囊组成。 这些胶囊是心型指向性，它们记录的信号通常称为“ Ambisonics A格式”。 然后，通过WXYZ通道的简单矩阵将A格式转换为B格式。

## Encoding (converting) mono, stereo or surround into B-format 将单声道，立体声或环绕声编码（转换）为B格式

B-format audio can also be **encoded**, or synthesized, out of regular audio recordings by an Ambisonics encoder.  
  
When encoding a mono track into B-format, you will need to decide where (in which direction) to position the mono signal in the 360-degree soundfield (the Waves [B360 Ambisonics Encoder](https://www.waves.com/plugins/b360-ambisonics-encoder) has panner-like controls which enable you to do that). The *output* of the encoding process will be a 4-channel B-format track, and the mono track will be present in each of these channels with the specific gain and phase that corresponds to its direction in the soundfield.  
  
Encoding multi-channel (stereo or surround) audio into Ambisonics follows the same principle. Each channel is encoded individually, like a mono track, in a set direction, and the results are summed together.  
  
Encoding regular audio into B-format is useful if you want to add sources to an existing B-format recording; mix a complete B-format track out of regular audio tracks by encoding each one separately and then combining them; or simply convert an entire multichannel mix into B-format.  
  
The Waves B360 plugin can address all the above use cases. It has mono, stereo, 5.1 and 7.1 components that encode the input onto B-format, with controls which allow you to position (pan) each element in the soundfield.

还可以通过Ambisonics编码器从常规音频记录中编码或合成B格式音频。

将单声道音轨编码为B格式时，您需要确定将单声道信号定位在360度声场中的位置（沿哪个方向）（Waves B360 Ambisonics Encoder具有类似于panner的控件，使您能够做到这一点） 。编码过程的输出将是一个4声道B格式音轨，单声道音轨将以与声音场方向相对应的特定增益和相位出现在每个这些声道中。

将多声道（立体声或环绕声）音频编码为Ambisonics遵循相同的原理。每个通道在设定的方向上像单声道一样分别编码，并将结果相加。

如果要将源添加到现有的B格式录音中，则将常规音频编码为B格式很有用。通过分别编码每个音频，然后将它们组合在一起，将完整的B格式音频与常规音频音频混合在一起；或简单地将整个多通道混音转换为B格式。

Waves B360插件可以解决以上所有用例。它具有单声道，立体声，5.1和7.1组件，可将输入编码为B格式，并带有控件，可让您在声场中定位（平移）每个元素。

## Playing back 播放

In principle, you can play back Ambisonics on almost any speaker array, recreating the spherical soundfield at the listening spot. But to do that, you need to **decode** the four B-format channels for the specific speaker array.  
  
Once again, decoding Ambisonics to speaker feeds is analogous to decoding M/S signals for stereo, only more complex. All four B-format channels are summed to each speaker feed. Each of the four channels is summed with different gain and phase, depending on the direction of the speaker. Some of the sources in the mix are summed in-phase while others are summed out-of-phase at each specific speaker. The result is that sources aligned with the direction of the speaker are louder, while those not aligned in the direction of the speaker are lower or cancel out. (Of course, if the speaker array is not fully spherical – for example, if it is just a regular stereo setup – the entire mix will be folded down when decoded to the available speakers.)

原则上，您可以在几乎任何扬声器阵列上播放Ambisonics，在聆听点重新创建球形声场。 但是要这样做，您需要解码特定扬声器阵列的四个B格式声道。

再一次，将Ambisonics解码为扬声器信号类似于对M / S信号进行立体声解码，只是更为复杂。 所有四个B格式声道都加到每个扬声器源中。 取决于扬声器的方向，四个声道中的每个声道的总增益和相位都不同。 在每个特定的扬声器中，混音中的某些信号源是同相求和的，而其他信号源是异相求和的。 结果是与扬声器方向对准的声源更大，而与扬声器方向对准的声源则更低或被抵消。 （当然，如果扬声器阵列不是完全球形的，例如，如果只是常规立体声设置，则将整个混音解码到可用扬声器时将被折叠。）

## Ambisonics on headphones? 在耳机上使用

Recently, Ambisonics has been adopted by the VR industry to deliver 360 audio for 360 videos, gaming and virtual reality experiences. Usually, the audio is experienced by the end user via headphones and a head-mounted display. This means that audio engineers who wish to hear the end result the way the user will hear it should monitor their Ambisonics mix on headphones.  
  
In addition, a multi-speaker spherical array for Ambisonics playback is highly expensive and often impractical even for professional studios.  
  
For both these reasons, it is advisable for audio engineers to be able to monitor their Ambisonics sessions on headphones.

最近，Ambisonics被VR行业所采用，可为360视频，游戏和虚拟现实体验提供360音频。 通常，最终用户会通过耳机和头戴式显示器来体验音频。 这意味着希望以最终用户听到声音的方式听到声音的音频工程师应该监听耳机上的Ambisonics混音。

此外，用于Ambisonics播放的多扬声器球形阵列非常昂贵，即使对于专业录音棚来说也常常是不切实际的。

由于这两个原因，建议音频工程师能够在耳机上监视其Ambisonics会话。

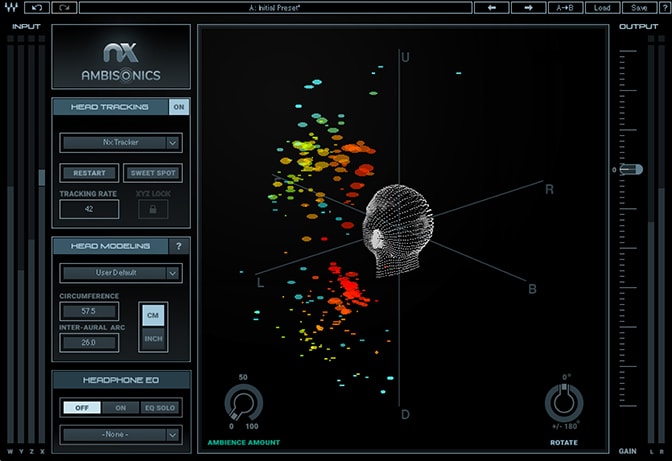
## How does this work? 他是怎么工作的

Spatial sound on headphone is made possible by **binaural audio technologies**. In essence, a binaural processor receives an audio input and a direction in which to position it. the processor adds auditory cues to the signal, so that when played back on headphones it is experienced at the set virtual position.  
  
The most common way to process Ambisonics for binaural spatial playback on headphones is to decode the Ambisonics channels for a certain speaker array – but then, instead of being sent to actual speakers, the feeds are sent to a binaural processor which virtually positions them at the direction that the actual speaker would have been. The result is that the immersive spherical soundfield is experienced by the listener when monitoring on headphones.  
  
The Waves [Nx Virtual Mix Room plugin](https://www.waves.com/plugins/nx) has a component called **Nx Ambisonics** which does just that. You feed Ambisonics channels into the plugin, and hear the soundfield reproduced on headphones, [complete with head tracking](https://www.waves.com/hardware/nx-head-tracker). (For convenience, all Waves Ambisonics tools – the B360 and Nx plugins, plus the Nx Head Tracker – are available together as part of the Waves [360° Ambisonics Tools](https://www.waves.com/hardware/360-ambisonics-tools) package.)

双耳音频技术使耳机上的空间声音成为可能。本质上，双耳处理器接收音频输入和放置音频的方向。处理器将听觉提示添加到信号中，以便在耳机上播放时可以在设置的虚拟位置进行体验。

处理Ambisonics以便在耳机上进行双耳空间回放的最常见方法是解码特定扬声器阵列的Ambisonics通道-但是，然后将提要发送到双耳处理器，而不是将其发送给实际的扬声器，双耳处理器实际上将它们定位在扬声器上。实际讲话者的方向。结果是当监听耳机时，听众会感受到沉浸式球形声场。

Waves Nx虚拟混音室插件具有一个名为Nx Ambisonics的组件，可以执行此操作。您将Ambisonics通道输入插件，并收听耳机再现的声场，并完成了头部跟踪。 （为方便起见，所有Waves Ambisonics工具-B360和Nx插件，以及Nx Head Tracker-作为Waves 360°Ambisonics工具包的一部分一起提供。）



Nx Ambisonics component of the Nx Virtual Mix Room plugin

Nx虚拟混合室插件的Nx Ambisonics组件

# Conclusion 结论

We’ve just scratched the surface of Ambisonics theory and practice, with the goal of offering audio engineers some understanding of the basic concepts and workflow of Ambisonics B-format. Future information will cover the finer points of mixing Ambisonics audio, so stay tuned!

我们只是对Ambisonics理论和实践进行了介绍，目的是使音频工程师对Ambisonics B格式的基本概念和工作流程有所了解。 未来的信息将涵盖混音Ambisonics音频的优点，敬请期待！