

Filed on behalf of: Meta Platforms, Inc.

Filed: July 16, 2024

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

META PLATFORMS, INC.
Petitioner,

v.

EIGHT KHZ, LLC
Patent Owner.

Case No. IPR2023-01003
U.S. Patent No. 9,226,090

**DECLARATION OF DR. CHRIS KYRIAKAKIS IN SUPPORT OF
PETITIONER'S REPLY TO PATENT OWNER'S RESPONSE**

I. INTRODUCTION

1. I have been retained by Meta Platforms, Inc. ("Petitioner") to provide my opinions concerning the validity of certain claims of U.S. Patent No. 9,226,090.

2. I have reviewed the Patent Owner's Response and Declaration of Dr. Durand Begault, as well as other materials and references cited in this Declaration.

II. CLAIM CONSTRUCTION

3. I understand that words of a patent claim are generally given their ordinary and customary meaning to a person of ordinary skill in the art at the time of the invention (POSITA), in the context of the claims, the written description, and the prosecution history. I understand that there are two exceptions to this rule: if the patentee disavows the full scope of the term either in the specification or during prosecution or if the patentee sets out his own definition and acts as his own lexicographer.

4. In paragraph 79 of Dr. Begault's declaration, he states: "The phrase 'away from' and 'proximate to' in limitation 8b. indicates a mode of binaural sound that would require externalization of the sound image. I have already established that the means of externalization for a listener requires both binaural sound with HRTFs and 6DOF tracking of the listener relative to the SLP that represents the talker." Dr. Begault further confirms that he believes that HRTFs and 6 degrees of freedom (6DOF) tracking is required based on the term "away from and proximate

to” in claims 1 and 15. Ex. 1051, 63:16-68:14. I disagree that the phrase “away from and proximate to” requires the use of HRTFs and 6DOF tracking. In my opinion, the term has its plain and ordinary meaning as it would have been understood by a POSITA in the context of the '090 Patent.

5. It is my opinion that a POSITA reading the claims in view of the specification would not have understood the '090 Patent to provide a special definition of “away from and proximate to” which would restrict the term to externalization of sound through HRTFs and 6DOF tracking of the listener to the sound localization point that represents the talker.

“away from and proximate to” (all claims)	
Petitioner's Interpretation	PO's Interpretation
plain and ordinary meaning of “a distance from”	externalization of sound through HRTFs and 6DOF tracking of the listener relative to the SLP that represents the talker (Begault Decl. ¶ 79 (cl. 8); <i>see also</i> ¶¶64-65 (cl. 1), ¶ 85 (cl. 15)).

A. The term has its plain meaning based on the intrinsic evidence.

6. The claims use the term according to its plain and ordinary meaning. The language of the claims recites “providing . . . the binaural sound to the listening person during the telephone call such that the origin of the binaural sound for the listening person occurs at the sound localization point in empty space that is away

from and proximate to the listening person.” *See* '090 Patent at claim 1; *see also* claim 8 and 15. A POSITA would have understood that localizing the sound “away from and proximate” would be performed through a number of different known methods, particularly since there is no language in the claims that restricts such localization to HRTFs.

7. For example, Dr. Begault admits interaural time difference (ITD) and interaural level difference (ILD) are not HRTFs. Ex. 1051, 23:6-7, 23:20-23; *see also* Ex. 1036 (“Within a 3-D sound system inclusion of ITD and IID based on the spectral alteration of the HRTF is considered to be more accurate and realistic than the simple IID and ITD differences described previously in the discussion on lateralization.”). But a POSITA would have understood that adjustments to ITD or ILD would have also been used to localize sound “away from and proximate to” the listening person, as confirmed by Dr. Begault’s own cited evidence. *See, e.g.*, Ex. 2026, 3-4 (“For regular audio playback (headphones, loudspeakers) the resulting *interaural time difference* (ITD) and *interaural level difference* (ILD) cues (i.e., the spatial cues at the level of the eardrums) only crudely approximate the cues evoked by sources that are physically placed at the auditory event positions.”) (cited by Begault Decl., ¶ 16).

8. Outside of ILDs, ITDs, and HRTFs, there were also many other known ways to provide a sound localization point that is “away from and proximate to” the

listener such that the sound localization point simulates an origin of the sound source. For example, Asahi (1979) discussed out-of-head externalization using stereophonic sound reproduction over headphones and without using HRTFs, but instead a dip filter. Ex. 1052. Similarly, Bauer (1961), which was also cited in Dr. Begault's book (Ex. 1026) at page 69), discussed substituting earphones for loudspeakers without substantial loss of space perspective using filters and delays, not HRTFs. Ex. 1053.

9. The dependent claims of the '090 Patent further confirm to a POSITA that localization "away from and proximate to" the listening person is not limited to adjusting HRTFs. For example, dependent claim 2 recites "adjusting an interaural time difference of the sound that the listening person hears," dependent claim 5 recites "adjusting an interaural level difference of the binaural sound that the listening person hears," and dependent claim 10 recites "adjusting head-related transfer functions (HRTFs) of the first individual to more accurately simulate real HRTFs of the first individual." Consequently, a POSITA reading the claims would have understood that localizing sound that is "away from and proximate to" a listener includes adjusting one or more of ITD, ILD, and/or HRTFs, since they are not the same.

10. Dependent claim 17 further confirms that adjusting ITD, ILD, and/or HRTFs are all options to localize sound that is "away from and proximate to" a

listener, specifically reciting “wherein the processor further executes the instructions to change an interaural time difference (ITD), an interaural level difference (ILD), and a head-related transfer function (HRTF) in order to position the sound localization point of the binaural voice to the location that is away from and proximate to the first person.”

11. The specification also identifies a number of known ways to provide binaural sound at a sound localization point that is “away from and proximate to” the listener outside of ILD, ITD, and HRTFs. For example, the specification explains that: “A location of sound and/or a direction to the sound can be altered or created using one or more of head-related transfer functions (HRTFs), accounting for differences in arrival times between the ears (interaural time differences), differences in amplitude or level of the sound between the ears (interaural level differences), asymmetrical spectral reflections from parts of the body (such as torso, shoulders, and pinnae), phase delays, group delays, and a ratio of the direct signal to the reverberated signal.” ’090 Patent, 3:59-67. Because the specification explains that “away from and proximate to” can be implemented using asymmetrical spectral reflections, phase delays, group delays, and a ratio of the direct signal to the reverberated signal, a POSITA would understand that the term is not limited to HRTFs.

12. Generally, the specification identifies a number of different aspects of sound that can be adjusted to create a sound localization point that is “away from and proximate to” the listener: “Adjustments are made to the sound for each of these uses since they are located at different locations around the table with respect to the sound localization points. These adjustments include, but are not limited to, calculations for and changes to ITDs, ILDs, HRTFs, attenuation, reverberations, and other aspects of sound discussed herein.” ’090 Patent, 35:49-55.

13. Further, the specification consistently explains that HRTFs are only an example of an aspect of sound that can be adjusted to create a sound localization point that is “away from and proximate to” the listener. The specification does not describe HRTFs as the only way to implement a sound localization point that is “away from and proximate to” the listener. For example, the specification states:

Adjustments to the sound are made so the sound emanates or originates from the sound location point. For example, the sound is processed so sound localization occurs for the listener to emanate from or originate at the sound localization point. For instance, the sound is processed to alter, add, or generate head-related transfer functions (HRTFs), interaural time differences, and/or interaural level differences to correspond to a fixed or variable distance from the listener, a coordinate location or a GPS location of the sound localization point.

'090 Patent, 4:20-29. A POSITA would understand that this passage expressly states that HRTFs are just one example method of creating a sound localization point that is "away from and proximate to" the listener.

14. In another passage, the specification states:

As the first individual moves with respect to the sound localization point, one or more aspects of the sound are altered to compensate for the movements such that the sound the first individual hears continues to be localized at the sound localization point. For example, one or more of a HRTF, an interaural level difference, and/or an interaural time difference are altered."

'090 Patent, 9:54-58. As I explained in my opening declaration, a POSITA would have known a number of different aspects of sound that can be altered to simulate the origin of sound that is "away from and proximate to" the listener. A POSITA would have understood that this passage of the patent is consistent with that knowledge, expressly stating that localization can occur by adjusting one or more aspects of sound.

15. A POSITA would have also understood that the "away from and proximate to" term does not require 6DOF tracking. The term "6DOF" or its variations (such as six degrees-of-freedom or 6 degrees of freedom) do not appear in the claims or '090 Patent itself.

16. A POSITA would have understood that 6DOF tracking is not necessary based on the plain language of the claims. First, the independent claims do not recite

tracking. Second, even if the claims did recite tracking, there is nothing in the claim language that limits the “away from and proximate to” term to 6DOF tracking. For example, claim 2 recites “tracking a head orientation,” but does not say 6DOF is required. Third, the independent claims do not use the term “tracking” even though dependent claims do use the term “tracking,” such as claim 2. Because of the use of “tracking” in some claims but not the independent claims, a POSITA would understand that the independent claims do not require “tracking.”

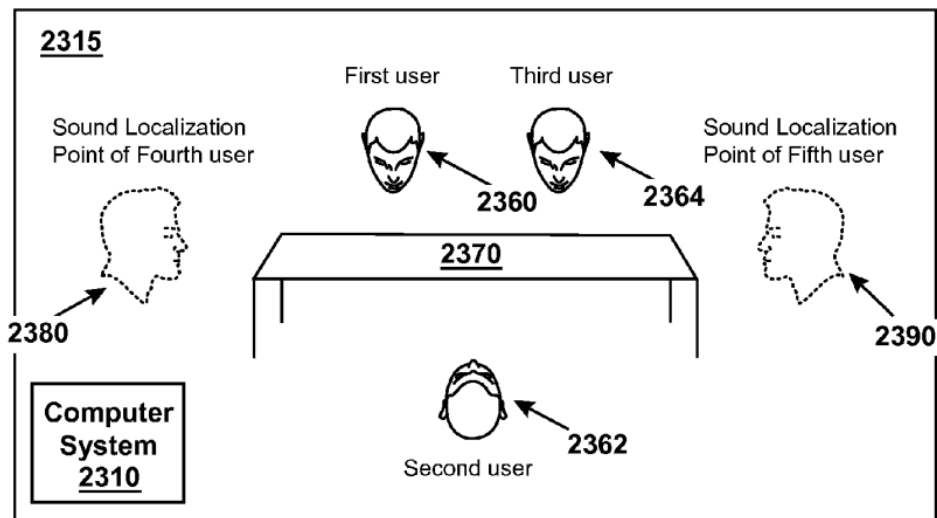
17. The specification also confirms that “away from and proximate to” does not require 6DOF tracking. The specification describes several embodiments that do not require 6DOF. For example, the specification describes an embodiment in which a listener has a telephone call with his mother in the kitchen, where the sound localization originates his mother's voice in an empty kitchen chair:

Consider an example in which a listener sits at his kitchen table and has an electronic telephone call with his mother. Both the listener and the mother wear electronic devices in or near their ears that capture, transmit, and provide sound for the telephone call. These electronic devices record binaural sound of the speaker and provide binaural sound to the listener. The listener designates an empty kitchen chair next to him as a sound localization point for the voice of his mother. During the telephone call, the voice of the mother appears to originate from the empty kitchen chair such that the listener localizes her voice to this location. From the point of view of the listener, the telephone conversation sounds like the mother is sitting in the kitchen chair and talking to the listener.

'090 Patent, 5:32-44.

18. A POSITA would understand that “away from and proximate to” covers this example because the empty kitchen chair is away from and yet still proximate to (i.e., a distance from) the listener. A POSITA would also understand that this example does not require tracking at all since there are no changes in movement or any direction that need to be tracked.

19. Similarly, the specification also describes a conference room embodiment with participants around a conference table. '090 Patent, 35:37-55. In this example, “the computer system elects to place the sound localization point 2390 of the fifth user 2368 at the empty chair.” '090 Patent, 35:42-44; Fig. 23.



'090 Patent, Figure 23 (excerpted)

20. Again, a POSITA would understand that “away from and proximate to” covers this example because the empty chair is away from and yet still proximate to

(i.e., a distance from) the listeners (first, second, and third users). Also again, a POSITA would understand that this example does not require tracking at all since there are no changes in any direction that need to be tracked. This is made even more clear because the specification later teaches that optionally the computer can base adjustments on head orientation or location changes—methods to track or otherwise reflect head orientation or location changes are not required. '090 Patent, 35:55-57 (“Further, these adjustments can occur based on movements to head orientation and location for each user.”).

21. Dr. Begault states in paragraph 55 of his declaration that “the specification provides many examples where 6DOF binaural synthesis is used to provide the voice of the talking person at a SLP away from and proximate from the listener at an externalized location.” Begault Decl., ¶ 55. But his statement (stating many examples, as opposed to all examples) confirms that the specification also describes embodiments without any 6DOF tracking, including the empty kitchen chair embodiment discussed above.

22. I note that Dr. Begault provides in paragraph 56 examples that he suggests confirm that 6DOF tracking is required. Ex. 2011, ¶ 56. But Dr. Begault never ties these examples to 6DOF tracking. Instead, he states that these examples require “intensive signal processing to dynamically update time-varying HRTF filters in response to listener’s location and head orientation.” *Id.*

23. Because the claim language does not restrict the “away from and proximate to” term to adjusting HRTFs and 6DOF tracking and the specification repeatedly provides examples of adjusting of other aspects of sound besides HRTFs, and without requiring tracking, a POSITA would understand the term to have its plain and ordinary meaning, which is not limited to HRTS and 6DOF tracking.

B. The extrinsic evidence does not establish that HRTFs and 6DOF tracking are required.

24. I note that Dr. Begault's declaration is confusing. Although Dr. Begault in paragraph 79, clearly states his opinion that “away from and proximate to” requires HRTFs and 6DOF tracking for claim 8, Dr. Begault appears to require HRTFs and 6DOF tracking for claim 1 based on the claim language “capturing, with an electronic earphone located at a head of a talking person, binaural sound that will be provided to a listening person.” Ex. 2011, ¶¶ 64-67. In particular, Dr. Begault states that “will be provided to a listening person” means that “the captured sound will be presented to the listener as binaural sound using binaural synthesis” (¶ 67) and that binaural synthesis requires both HRTFs and 6DOF tracking (¶¶ 64-65). *See also* Ex. 2011, ¶ 15 (“Stereo audio is based on techniques for loudspeaker playback as opposed to headphones, and does not include the localization cues used in binaural synthesis (HRTFs, 6DOF tracking).”).

25. For claim 15, Dr. Begault appears to restrict the claim scope to binaural synthesis requiring HRTFs and 6DOFs based on two limitations: “wearable

electronic device (§ 59) and “adjust the binaural voice of the second person during the electronic phone call in order to position a sound localization point of the binaural voice to a location that is away from and proximate to the first person” (§ 86). I disagree that either limitation should be so limited.

26. Dr. Begault clarified during deposition that his analysis of “away from and proximate to” in claim 8 applies to both claim 1 and 15. Ex. 1051. But to be clear, it is my opinion that a POSITA reading claims 1 and 15 would not require HRTFs or 6DOF tracking based on any language in those claims. Instead, it is my opinion that a POSITA would give those terms their plain and ordinary meaning.

27. I note that claim 8 does require that “the origin of the voice of the second individual appears to the first individual to originate from a fixed location that is the sound localization point in empty space while the first individual *moves and changes a location and a head orientation* with respect to the sound localization point.” As I mentioned above, there is no language in the claims that would preclude other types of tracking or other methods of reflecting changes in location or head orientation.

28. Claim 8 further requires processing outside of HRTFs and 6DOF tracking, reciting in limitation 8[d] “alter[ing] an interaural time difference [ITD]” when the head orientation of the first listener moves. But as Dr. Begault confirmed, an HRTF includes ITD (and other sound localization cues), and is not itself the same

as ITD. Ex. 1051, 23:6-7, 23:20-23. Further, the claim does not use the term “tracking” even though other dependent claims do use the term “tracking,” such as claim 2.

29. Throughout his declaration, Dr. Begault states that externalization and binaural synthesis requires HRTFs and 6DOF tracking. *See* Ex. 2011, ¶ 15 (“Stereo audio is based on techniques for loudspeaker playback as opposed to headphones, and does not include the localization cues used in binaural synthesis (HRTFs, 6DOF tracking).”); ¶ 64 (“The means of providing binaural sound of the talker is to use binaural synthesis techniques that include 6DOF tracking such that externalization of the sound source is ensured.”); ¶ 79 (“I have already established that the means of externalization for a listener requires both binaural sound with HRTFs and 6DOF tracking of the listener relative to the SLP that represents the talker.”). I disagree.

30. It is my opinion that externalization and binaural synthesis do not require HRTFs or 6DOF tracking. Instead, as the '090 Patent expressly explains: “Binaural synthesis transforms a sound source with no positional information to a virtual sound source with respect to a head of a listener.” '090 Patent, 28:45-47. Based on the discussion in the '090 Patent, binaural synthesis can encompass any method that creates a virtual sound source with respect to a head of a listener. The '090 Patent goes on to provide an example: “By way of example, binaural synthesis involves convolving a mono sound with HRIRs to generate a synthetic binaural

signal that includes directions information of the sound source.” ’090 Patent, 28:56-59. Dr. Begault agreed that these were both ways of describing binaural synthesis. Ex. 2051, 43:22-44:3, 44:11-17. Binaural synthesis synthesizes binaural sound. There is no restriction as to how that binaural sound is synthesized.

C. Dr. Begault’s opinion that “away from and proximate to” requires HRTFs is incorrect.

31. In paragraph 86 of his declaration, Dr. Begault states: “In order to ‘adjust the binaural voice of the second person during the electronic phone call in order to position a sound localization point of the binaural voice to a location that is away from and proximate to the first person,’ one would necessarily require binaural synthesis and a processor to process the voice of the talker with HRTFs.” Dr. Begault also repeatedly states that “away from and proximate to” means externalization, which necessarily requires HRTFs. I disagree that HRTFs are required. In particular, externalization can be implemented in a number of different methods. For example, as I discussed above, the specification provides numerous examples of aspects of sound that can be adjusted to localize sound at a point that is “away from and proximate to the listener.”

32. Dr. Begault does not explain why adjusting one or more aspects of sound expressly identified in the patent, such as ITD, ILD, attenuation, group delay, phase delay, or reverberations, does not localize the sound to a location that is “away from and proximate to” the listener, as the ’090 Patent teaches.

33. In fact, Dr. Begault has previously acknowledged adjustments to aspects of sound can localize that sound at a point “away from and proximate to” the listener, and explained in his book, “the same physical cue—reverberation—is key for the effective simulation of distance and the perception of an environmental context.” Ex. 1036 at 79. This is consistent with the discussions in the patent that explain that localizing sound “away and proximate to” the listener can simply include adjusting reverberation:

By way of example, one or more adjustments are made to a HRTF, an interaural time difference, and an interaural level difference of the sound provided to the listener. Sound can also be adjusted based on calculations made for sound attenuation and sound reverberation.

'090 Patent, 10:46-50.

34. Further, Dr. Begault himself confirmed that externalization does not require HRTFs and 6DOF tracking. For example, Dr. Begault confirmed that externalization can be accomplished through stereo processing over loudspeakers. Ex. 1051, 68:23-69:8, 70:8-10.

35. Although Dr. Begault qualifies his statement that externalization over headphones (as opposed to loudspeakers) must be accomplished through HRTFs, there are numerous other known ways to implement externalization over headphones. For example, as I discussed above, Asahi (1979) discussed out-of-head externalization using stereophonic sound reproduction over headphones and without

using HRTFs, but instead a dip filter. Ex. 1052. Similarly, Bauer (1961), which was also cited in Dr. Begault's book (Ex. 1026) at page 69), discussed substituting earphones for loudspeakers without substantial loss of space perspective using filters and delays, not HRTFs. Ex. 1053.

D. Dr. Begault's opinion that "away from and proximate to" requires 6DOF tracking

36. In paragraph 64 of his declaration, in relation to claim 1, Dr. Begault states that: "The means of providing the binaural sound of the talker is to use binaural synthesis techniques that include 6DOF tracking such that externalization of the sound source is ensured." I disagree that 6DOF tracking is required.

37. In particular, claim 1 only requires that the sound is localized to originate "away from and proximate to" the listener. No tracking is required at all. Accordingly, to localize the sound to, for example, an empty space to the left of the listener, the ITD could be adjusted such that the left ear receives the sound of the talker slightly before the right ear. Similarly, the ILD could be adjusted so that the left ear perceives sound slightly louder than the right ear. 6DOF tracking, or any tracking at all, would not be required to make that adjustment. Instead, an empty space to the left of the speaker could simply be designated without any tracking.

38. Like claim 1, claim 15 only requires that the sound is localized to originate "away from and proximate to" the listener. No tracking is required at all. Again, to localize the sound to, for example, an empty space to the left of the listener,

the ITD could be adjusted such that the left ear receives the sound of the talker slightly before the right ear. Similarly, the ILD could be adjusted so that the left ear perceives sound slightly louder than the right ear. 6DOF tracking, or any tracking at all, would not be required to make that adjustment. Instead, an empty space to the left of the speaker could simply be designated by the processor without any tracking.

III. DR. BEGAULT'S MISCHARACTERIZATION OF THE PRIOR ART

A. Andrea reflects the same disclosure in the '090 Patent.

39. In paragraph 76 of his declaration, Dr. Begault states that “if one were to add Andrea’s noise cancellation microphone system to Bedingfield’s headphones, one would cancel or minimize surrounding sound. This is the opposite of the goal of Claim 1, which is to capture the voice of the talker *and* surrounding sounds in the environment.” I disagree. It is my opinion that Andrea’s disclosure on capturing binaural sound is effectively the same as the disclosure in the '090 Patent.

40. The '090 Patent teaches capturing binaural sound may include capturing background or other environmental sounds. Contrary to Dr. Begault’s statement that cancelling or minimizing surrounding sound is not a goal, the '090 Patent teaches that such surrounding sound can be removed. '090 Patent, 25:39-41 (“Ambient noise and attenuation due to environmental conditions can be filtered.”). In fact, the '090 Patent recognizes that the use of multiple microphones can be used to isolate the vocal and non-vocal components of sound. '090 Patent, 39:65-40:3

(“The vocal sound captured from the third microphone is used to enhance the sound captured by the fourth microphone, to isolate and cancel non-vocal components of sound, and to improve the clarity of a vocal part of the signal that is ultimately provided to the listener.”).

41. Like the '090 Patent, Andrea teaches “an audio receiving means on each earbud” that captures sound perceived by the user, including “unwanted sound perceived by the user.” Ex. 1007, 14:55-61, 15:34-37. Andrea states “the sound picked up by the audio receiving means on the earbuds may be converted to audio signals that are transmitted to an external device comprising a processor.” Ex. 1007, 16:21-24. Once two distinct audio signals have been captured by the audio receiving means on each earbud, Andrea teaches that the processor, then optionally, “may execute executable instructions causing it to generate an audio signal that substantially fails to reflect noise generated” by applying beamforming algorithms. Ex. 1007, 16:24-28. Also, like the '090 Patent, Andrea recognizes that the two audio signals can be used to isolate the vocal and non-vocal components of the sound using audio enhancement techniques like, ANR. Ex. 1007, 15:16-18.

B. Bedingfield is a binaural reference.

42. The Patent Owner's Response states that Bedingfield has “nothing to do with binaural sound or binaural synthesis.” POR at 58-59. I disagree. Bedingfield discloses providing binaural sound.

43. I note that Dr. Begault does not state in his declaration that Bedingfield does not relate to binaural sound. Instead, Dr. Begault only states that Bedingfield “(1) never mentions the term HRTF...(2) never mentions altering frequencies to correspond to the HRTF spectrum...(3) never mentions sound source externalization.” Begault, ¶ 43. Although Bedingfield does not use the terms HRTF or externalization, it is my opinion that both are disclosed.

44. Patent Owner states that the term “binaural” “has two parts: ‘bi’ meaning ‘two,’ and ‘aural’ meaning ‘of or relating to the ear or to the sense of hearing.’” Ex. 1027 at 18. I agree, and Bedingfield discloses providing binaural sound because Bedingfield teaches providing different sounds to each ear.

45. In particular, Bedingfield recognizes “the human hearing mechanism (including the ears, nerves and brain processing, subsequently referred to as ‘the ear’) detects placement of a speaker (or source of sound)...by detecting [differences]...perceived by the listener’s two ears.” Ex. 1005, 3:35-39. Based on that principle, Bedingfield teaches adjusting differences in the audio signal provided to the left ear as compared to the audio signal provided to the right ear to localize the sound source so that it appears to originate from different locations. Ex. 1005, 8:35-36 (“The SSAPC may be configured to arrange the participants 404a-404d in a synthetic audio placement, using the left and right channels to provide stereo separation.”); 9:50-57 (“[W]hen participant 604 is speaking, the listener 602 may

hear 67% of the speaker's sound in the listener's 602 left ear and 33% of the speaker's sound in the listener's 602 right ear, thus creating a balance effect to aurally locate the participant 604 to the left of the listener 602.”).

46. This is seen in Figures 14 and 15, where Bedingfield shows different volume levels in the left channel and right channel depending where the sound appears to originate in relation to the listener.

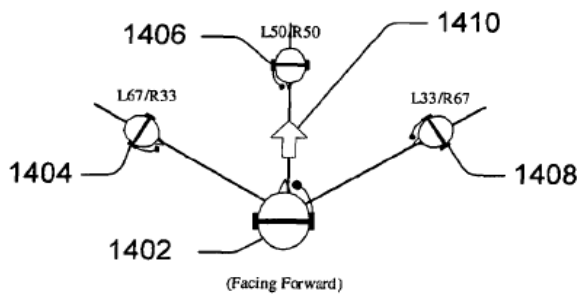


Fig. 14

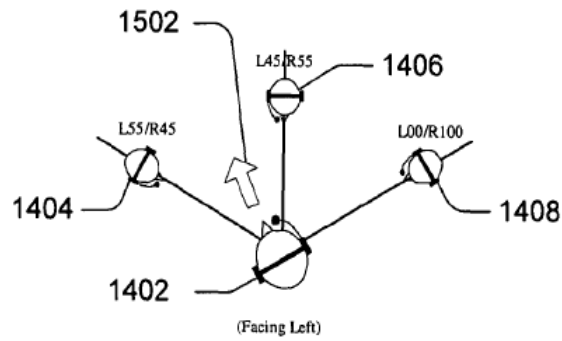


Fig. 15

47. Similarly, as Bedingfield explains, in a sound field that includes four participants configured in a circular configuration, a “listener may hear a second speaker, located at the 3 o’clock position relative to the listener, in stereo sound with an increased volume in the right audio channel and a decreased volume in the left audio channel, compared to the right and left volume levels received from the first speaker.” Ex. 1005, 3:62-4:12, Fig.19a.

48. In Paragraph 83 of declaration, Dr. Begault states that: “Bedingfield’s focus, volume, and change of location (Fig. 3, block 322) would only accomplish a change in the inside-the-head location of a source.” I disagree.

49. Bedingfield teaches that the volume change will localize the sound behind the listener. Ex. 1005, 13:30-43 (in FIG. 15, the participant 1408 that is now speaking will sound as if the speaker were *located behind* the listener 1402.”). A POSITA reading Bedingfield would have understood that perceiving the speaker behind the listener is externalized and not inside the head. For example, Dr. Begault’s own cited evidence states that adjustment to ILD (which encompasses change to volume) will approximate the cues by sources “that are physically placed at auditory event positions.” *See, e.g.*, Ex. 2026, 3-4 (“For regular audio playback (headphones, loudspeakers) the resulting *interaural time difference* (ITD) and *interaural level difference* (ILD) cues (i.e., the spatial cues at the level of the eardrums) only crudely approximate the cues evoked by sources that are physically placed at the auditory event positions.”) (cited by Begault Decl., ¶ 16).

50. Through these examples and discussion, Bedingfield teaches a system that simulates auditory cues containing natural acoustic differences between the two ears. Thus, a POSITA would have understood that Bedingfield teaches binaural sound.

C. A POSITA would have looked at Ballas for a headphone system.

51. In paragraph 104 of his declaration, Dr. Begault states that: “A POSITA interested in binaural recording or synthesis has no motivation looking at loudspeaker-based system disclosed in Ballas and combine it with a system based exclusively on earphones.” I disagree.

52. Ballas discloses a transaural system. As Kaiser (2011) explains, “Transaural audio is a three-dimensional sound spatialization technique which is capable of reproducing binaural signals over loudspeakers.” Ex. 1054 at 32. Transaural sound relates to binaural sound since it is binaural sound processed and delivered through speakers instead of headphones.

53. Further, a POSITA interested in binaural recording or synthesis would have looked at Ballas because Ballas itself states that it relates to providing binaural sound. In particular, Ballas teaches that the “speakers 110 [that] are attached close to the temple of a user’s head” provide binaural sound. For example, Ballas expressly states “the audio signals generated by the speakers is binaural so that the audio signals from right and left speakers are delivered to right and left ears of the user, respectively.” Ex. 1008, claim 2. Similarly, Ballas teaches that the “spatialized audio signals are amplified” and the “amplified signals are binaural in nature (i.e., left channels signals are supplied to the left ear and right channel signals are supplied to the right ear[)].” Ex. 1008, ¶ 25.

54. In paragraph 105 of his declaration, Dr. Begault states: “A POSITA would be immediately concerned with the obvious cross-talk that would occur between the two external loudspeaker, and the lack of a solution to the problem in Ballas.” But there were well-known methods to adjust filters from headphones to speakers, and from speakers to headphones.

55. For example, the '090 Patent and Dr. Begault discussed the use of one method—crosstalk cancellation—to adjust filters for headphones to speakers. The '090 Patent states:

Consider an example in which a listener is located in a room with speakers, and the listener does not wear headphones or an electronic device providing sound to the ears. The speakers provide binaural sound to the listener, and a crosstalk cancellation (CTC) system adjusts, removes, filters, or moves a crosstalk location in real-time as the listener moves around the room.

'090 Patent, 8:44-50. Dr. Begault also discusses crosstalk cancellation in his book: “Headphone listening conditions can be roughly approximated from stereo loudspeakers using a technique called cross-talk cancellation.” Ex. 1036, 26, 184-87. And the technique is also discussed in one of Dr. Begault's cited references. Ex. 2026, 4.

56. Further, there were other known methods of adjusting speaker signals so they can play binaurally over earphones. For example, Bauer (1961) discusses one known method involving filters and delays without using HRTFs. Ex. 1054.

57. Thus, a POSITA designing a spatialized headphone system would have looked at known methods used in spatialized speaker systems, and the other way around, including Ballas.

Executed on July 16, 2024 in Los Angeles, California.



Dr. Chris Kyriakakis