

UNITED STATES PATENT AND TRADEMARK OFFICE

In re *Inter Partes* Review of:)
U.S. Patent No. 8,503,691)
Issued: August 6, 2013)
Application No.: 12/139,333)
Filing Date: June 13, 2008)

**For: Virtual Microphone Arrays Using Dual Omnidirectional Microphone
Array (DOMA)**

FILED VIA P-TACTS

**DECLARATION OF DR. MICHAEL JOHNSON IN SUPPORT OF
PETITION FOR *INTER PARTES* REVIEW OF
U.S. PATENT NO. 8,503,691**

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

TABLE OF CONTENTS

I.	INTRODUCTION -----	1
II.	EXPERIENCE AND QUALIFICATIONS -----	2
III.	UNDERSTANDING OF RELEVANT LEGAL PRINCIPLES -----	5
	A. Claim Construction -----	5
	B. Obviousness -----	6
IV.	MATERIALS REVIEWED AND CONSIDERED -----	8
V.	LEVEL OF SKILL IN THE ART AND PERSPECTIVE APPLIED IN THIS DECLARATION -----	10
VI.	BACKGROUND -----	11
	A. In 1982, Griffiths and Jim Publish Their Seminal Article on the GSC. -----	12
	B. The GSC Was Applied to Microphone Arrays to Reduce Noise -----	15
	C. A 2001 Book Edited by Brandstein Illustrates How to Apply a GSC to a Microphone Arrays to Reduce Noise -----	16
	D. Near-field Compensation Techniques Were Well Known -----	23
VII.	THE '691 PATENT -----	25
	A. Specification -----	25
	B. Prosecution History -----	29
	C. The Earliest Possible Priority Date of the '691 Patent Claims Is June 13, 2007. -----	31

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

D.	The References Are Prior Art.	31
E.	The Asserted References Are Analogous Art.	31
VIII.	GROUND OF UNPATENTABILITY	32
A.	Ground 1: Brandstein Renders Obvious Claims 1-46	32
1.	Motivation to Combine	32
2.	Reasonable Expectation of Success	35
3.	Claim 1	37
a)	[1pre]: “A microphone array comprising:”	37
b)	[1a]: “a first virtual microphone comprising a first combination of a first microphone signal and a second microphone signal,”	38
c)	[1b]: “the first virtual microphone having a first linear response to speech and a first linear response to noise,”	40
d)	[1c]: “the first linear response to speech being substantially similar across a plurality of frequencies for a speech source located at an angle relative to an axis of the microphone array,”	44
e)	[1d]: “wherein the first microphone signal is generated by a first physical microphone and the second microphone signal is generated by a second physical microphone; and”	46
f)	[1e]: “a second virtual microphone comprising a second combination of the first microphone signal and the second microphone signal,”	47

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

- g) [1f]: “the second virtual microphone having a second linear response to speech and a second linear response to noise,” ----- 49
- h) [1g]: “the second linear response to noise being substantially similar to the first linear response to noise, and the second linear response to speech being substantially dissimilar to the first linear response to speech,” ----- 50
- i) [1h]: “wherein at least one of the first linear response to noise and the second linear response to noise is being non-zero in a direction towards a source of noise.” ----- 54
- 4. Claim 2: “The microphone array of claim 1, wherein the first and second physical microphones are omnidirectional.” ----- 54
- 5. Claim 3: “The microphone array of claim 1, wherein the first linear response to speech is devoid of a null, wherein the speech is human speech.” ----- 55
- 6. Claim 4: “The microphone array of claim 3, wherein the second linear response to speech of the second virtual microphone includes a single null oriented in a direction toward a source of the speech.” ----- 56
- 7. Claim 5: “The microphone array of claim 4, wherein the single null is a region of the second linear response to speech having a measured response level that is lower than the measured response level of any other region of the second linear response to speech.” ----- 60
- 8. Claim 6: “The microphone array of claim 4, wherein the second linear response to speech includes a primary lobe oriented in a direction away from the source of the speech.” ----- 60

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

9. Claim 7: “The microphone array of claim 6, wherein the primary lobe is a region of the second linear response to speech having a measured response level that is greater than the measured response level of any other region of the second linear response to speech.” ----- 62
10. Claim 8: “The microphone array of claim 4, wherein the first physical microphone and the second physical microphone are positioned along an axis and separated by a first distance.” ----- 62
11. Claim 9: “The microphone array of claim 8, wherein a midpoint of the axis is a second distance from the speech source that generates the speech, wherein the speech source is located in a direction defined by an angle relative to the midpoint.” ----- 65
12. Claim 10: “The microphone array of claim 9, wherein the first virtual microphone comprises the second microphone signal subtracted from the first microphone signal.” ----- 66
13. Claim 11: “The microphone array of claim 10, wherein the first microphone signal is delayed.” ----- 68
14. Claim 12: “The microphone array of claim 11, wherein the delay is raised to a power that is proportional to a time difference between arrival of the speech at the first virtual microphone and arrival of the speech at the second virtual microphone.” ----- 72
15. Claim 13: “The microphone array of claim 11, wherein the delay is raised to a power that is proportional to a sampling frequency multiplied by a quantity equal to a third distance subtracted from a fourth distance, the third distance being between the first physical microphone and the speech source and

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

the fourth distance being between the second physical
microphone and the speech source.”----- 77

16. Claim 14: “The microphone array of claim 10,
wherein the second microphone signal is multiplied
by a ratio, wherein the ratio is a ratio of a third
distance to a fourth distance, the third distance being
between the first physical microphone and the speech
source and the fourth distance being between the
second physical microphone and the speech source.” ----- 81
17. Claim 15: “The microphone array of claim 9, wherein
the second virtual microphone comprises the first
microphone signal subtracted from the second
microphone signal.”----- 82
18. Claim 16: “The microphone array of claim 15,
wherein the first microphone signal is delayed.”----- 83
19. Claim 17: “The microphone array of claim 16,
wherein the delay is raised to a power that is
proportional to a time difference between arrival of
the speech at the first virtual microphone and arrival
of the speech at the second virtual microphone.” ----- 84
20. Claim 18: “The microphone array of claim 16,
wherein the power is proportional to a sampling
frequency multiplied by a quantity equal to a third
distance subtracted from a fourth distance, the third
distance being between the first physical microphone
and the speech source and the fourth distance being
between the second physical microphone and the
speech source.” ----- 84
21. Claim 19: “The microphone array of claim 18,
wherein the first microphone signal is multiplied by a
ratio, wherein the ratio is a ratio of the third distance
to the fourth distance.”----- 84

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

22. Claim 20: “The microphone array of claim 4, wherein the single null is located at a distance from at least one of the first physical microphone and the second physical microphone where the source of the speech is expected to be.” ----- 85
23. Claim 21: “The microphone array of claim 1, wherein the first virtual microphone comprises the second microphone signal subtracted from a delayed version of the first microphone signal.”----- 85
24. Claim 22: “The microphone array of claim 21, wherein the second virtual microphone comprises a delayed version of the first microphone signal subtracted from the second microphone signal.”----- 86
25. Claim 23 ----- 88
- a) [23pre]: “A microphone array comprising:” ----- 88
- b) [23a]: “a first virtual microphone formed from a first combination of a first microphone signal and a second microphone signal,” ----- 88
- c) [23b]: “wherein the first microphone signal is generated by a first omnidirectional microphone and the second microphone signal is generated by a second omnidirectional microphone; and” ----- 88
- d) [23c]: “a second virtual microphone formed from a second combination of the first microphone signal and the second microphone signal,” ----- 88
- e) [23d]: “wherein the second combination is different from the first combination;”----- 89

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

- f) [23e]: “wherein the first virtual microphone has a first linear response to speech that is substantially similar across a plurality of frequencies for a speech source within a predetermined angle relative to an axis of the microphone array and devoid of a null, and a first linear response to noise that is devoid of a null,” ----- 89
 - g) [23f]: “wherein the second virtual microphone has a second linear response to speech that has a single null oriented in a direction toward a source of the speech and a second linear response to noise that is devoid of a null,” ----- 90
 - h) [23g]: “wherein the second linear response to noise is substantially similar to the first linear response to noise and the second linear response to speech is substantially dissimilar to the first linear response to speech,” ----- 91
 - i) [23h]: “wherein the speech is human speech.” ----- 92
26. Claim 24: “The microphone array of claim 23, wherein the single null is a region of the second linear response to speech having a measured response level that is lower than the measured response level of any other region of the second linear response to speech.” ----- 92
27. Claim 25: “The microphone array of claim 23, wherein the second linear response to speech includes a primary lobe oriented in a direction away from the source of the speech.” ----- 92
28. Claim 26: “The microphone array of claim 25, wherein the primary lobe is a region of the second linear response to speech having a measured response

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

	level that is greater than the measured response level of any other region of the second linear response to speech.” -----	93
29.	Claim 27 -----	93
a)	[27pre]: “A device comprising:” -----	93
b)	[27a]: “a first microphone outputting a first microphone signal and a second microphone outputting a second microphone signal; and” -----	93
c)	[27b]: “a processing component coupled to the first microphone signal and the second microphone signal,” -----	93
d)	[27c]: “the processing component generating a virtual microphone array comprising a first virtual microphone and a second virtual microphone,” -----	94
e)	[27d]: “wherein the first virtual microphone comprises a first combination of the first microphone signal and the second microphone signal,” -----	94
f)	[27e]: “wherein the second virtual microphone comprises a second combination of the first microphone signal and the second microphone signal,” -----	95
g)	[27f]: “wherein the second combination is different from the first combination,” -----	95
h)	[27g]: “the first virtual microphone having a first linear response to speech and a first linear response to noise, the first linear response to speech being substantially similar across a	

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

	plurality of frequencies for a speech source at an angle relative to an axis of the microphone array,” -----	95
i)	[27h]: “the second virtual microphone having a second linear response to speech and a second linear response to noise,” -----	95
j)	[27i]: “the second linear response to noise being substantially similar to the first linear response to noise, and the second linear response to speech being substantially dissimilar to the first linear response to speech,” -----	95
k)	[27j]: “wherein one of the first linear response to speech and the second linear response to speech is reduced in a direction towards a source of speech.” -----	96
30.	Claim 28 -----	96
a)	[28pre] “A device comprising:” -----	96
b)	[28a]: “a first microphone outputting a first microphone signal and a second microphone outputting a second microphone signal, wherein the first microphone and the second microphone are omnidirectional microphones; and” -----	96
c)	[28b]: “a virtual microphone array comprising a first virtual microphone and a second virtual microphone,” -----	97
d)	[28c]: “the first virtual microphone having a first linear response to speech based on distances from a first physical microphone and	

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

- a second physical microphone to a source of
speech, and ----- 97
- e) [28d] a first linear response to noise, the first
linear response to speech being substantially
similar across a plurality of frequencies for a
speech source within a predetermined angle
relative to an axis of the microphone array,”----- 98
- f) [28e]: “the second virtual microphone having a
second linear response to speech and a second
linear response to noise,” ----- 98
- g) [28f]: “the second linear response to noise
being substantially the first linear response to
noise, and the second linear response to speech
being substantially dissimilar to the first linear
response to speech,” ----- 98
- h) [28g]: “wherein the distances form a ratio with
which to direct the first and second linear
responses to speech toward the source of
speech,” ----- 99
- i) [28h]: “wherein the first virtual microphone
comprises a first combination of the first
microphone signal and the second microphone
signal,” ----- 99
- j) [28i]: “wherein the second virtual microphone
comprises a second combination of the first
microphone signal and the second microphone
signal,” ----- 99
- k) [28j]: “wherein the second combination is
different from the first combination,”----- 100

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

l)	[28k]: “wherein the first virtual microphone and the second virtual microphone are distinct virtual directional microphones.”-----	100
31.	Claim 29 -----	100
a)	[29pre]: “A device comprising:”-----	100
b)	[29a]: “a first physical microphone generating a first microphone signal;” -----	100
c)	[29b]: “a second physical microphone generating a second microphone signal; and” -----	101
d)	[29c]: “a processing component coupled to the first microphone signal and the second microphone signal,” -----	101
e)	[29d]: “the processing component generating a virtual microphone array comprising a first virtual microphone and a second virtual microphone,” -----	101
f)	[29e]: “the first virtual microphone having a first linear response to speech and a first linear response to noise,”-----	101
g)	[29f]: “the first linear response to speech being substantially similar across a plurality of frequencies for a speech source at an angle relative to an axis of the microphone array,”-----	101
h)	[29g]: “the second virtual microphone having a second linear response to speech and a second linear response to noise,” -----	102
i)	[29h]: “the second linear response to noise being substantially similar to the first linear	

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

- response to noise, and the second linear
response to speech being substantially
dissimilar to the first linear response to
speech,” ----- 102
- j) [29i]: “wherein at least one of the first linear
response to noise and the second linear
response to noise is non-zero in a direction
towards a source of noise,” ----- 102
- k) [29j]: “wherein the first virtual microphone
comprises the second microphone signal
subtracted from a delayed version of the first
microphone signal,” ----- 102
- l) [29k]: “wherein the second virtual microphone
comprises a delayed version of the first
microphone signal subtracted from the second
microphone signal.” ----- 103
32. Claim 30: “The device of claim 29, wherein the first
linear response to speech is devoid of a null, wherein
the speech is human speech.” ----- 103
33. Claim 31: “The device of claim 30, wherein the
second linear response to speech of the second virtual
microphone includes a single null oriented in a
direction toward a source of the speech.” ----- 103
34. Claim 32: “The device of claim 31, wherein the
single null is a region of the second linear response to
speech having a measured response level that is
lower than the measured response level of any other
region of the second linear response to speech.” ----- 103
35. Claim 33: “The device of claim 31, wherein the
second linear response to speech includes a primary

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

- lobe oriented in a direction away from the source of
the speech.” ----- 104
36. Claim 34: “The device of claim 33, wherein the
primary lobe is a region of the second linear response
to speech having a measured response level that is
greater than the measured response level of any other
region of the second linear response to speech.” ----- 104
37. Claim 35: “The device of claim 31, wherein the first
physical microphone and the second physical
microphone are positioned along an axis and
separated by a first distance.” ----- 104
38. Claim 36: “The device of claim 35, wherein a
midpoint of the axis is a second distance from the
speech source that generates the speech, wherein the
speech source is located in a direction defined by an
angle relative to the midpoint.” ----- 104
39. Claim 37: “The device of claim 36, wherein one or
more of the first microphone signal and the second
microphone signal is delayed.” ----- 105
40. Claim 38: “The device of claim 37, wherein the delay
is raised to a power that is proportional to a time
difference between arrival of the speech at the first
virtual microphone and arrival of the speech at the
second virtual microphone.” ----- 105
41. Claim 39: “The device of claim 38, wherein the
power is proportional to a sampling frequency
multiplied by a quantity equal to a third distance
subtracted from a fourth distance, the third distance
being between the first physical microphone and the
speech source and the fourth distance being between
the second physical microphone and the speech
source.” ----- 105

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

42.	Claim 40: “The device of claim 36, wherein one or more of the first microphone signal and the second microphone signal is multiplied by a gain factor.”-----	105
43.	Claim 41 -----	106
a)	[41pre]: “A sensor comprising:”-----	106
b)	[41a]: “a physical microphone array including a first physical microphone and a second physical microphone,” -----	106
c)	[41b]: “the first physical microphone outputting a first microphone signal and the second physical microphone outputting a second microphone signal; and” -----	107
d)	[41c]: “a virtual microphone array comprising a first virtual microphone and a second virtual microphone,” -----	107
e)	[41d]: “the first virtual microphone comprising a first combination of the first microphone signal and the second microphone signal and” -----	107
f)	[41e]: “having a first linear response to speech and a first linear response to noise, the first linear response to speech being substantially similar across a plurality of frequencies for a source of speech at an angle relative to an axis of the microphone array,” -----	107
g)	[41f]: “the second virtual microphone comprising a second combination of the first microphone signal and the second microphone signal and”-----	108

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

- h) [41g]: “having a second linear response to speech and a second linear response to noise,” ----- 108
- i) [41h]: “the second linear response to noise being substantially similar to the first linear response to noise and the second linear response to speech being substantially dissimilar to the first linear response to speech,” ----- 108
- j) [41i]: “wherein one of the first linear response to speech and the second linear response to speech is reduced in a direction towards a source of speech,” ----- 108
- k) [41j]: “wherein the virtual microphone array includes a single null oriented in a direction toward the source of speech of a human speaker.” ----- 109
- 44. Claim 42: “The sensor of claim 41, wherein the first linear response to speech is devoid of a null, wherein the second linear response to speech of the second virtual microphone includes the single null.”----- 109
- 45. Claim 43: “The sensor of claim 42, wherein the single null is a region of the second linear response to speech having a measured response level that is lower than the measured response level of any other region of the second linear response to speech.”----- 109
- 46. Claim 44: “The sensor of claim 42, wherein the second linear response to speech includes a primary lobe oriented in a direction away from the source of the speech.” ----- 109
- 47. Claim 45: “The sensor of claim 44, wherein the primary lobe is a region of the second linear response

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

	to speech having a measured response level that is greater than the measured response level of any other region of the second linear response to speech.” -----	110
48.	Claim 46: “The sensor of claim 41, wherein the single null is located at a distance from the physical microphone array where the source of the speech is expected to be.” -----	110
B.	Ground 2: Brandstein in combination with McCowan -----	110
1.	Motivation to Combine -----	110
2.	Expectation of Success -----	114
3.	Claim 12 -----	114
4.	Claim 13 -----	115
5.	Claim 14 -----	116
6.	Claim 17 -----	116
7.	Claim 18 -----	117
8.	Claim 19 -----	117
9.	Claim 28 -----	117
10.	Claim 38 -----	117
11.	Claim 39 -----	118
IX.	CONCLUSION -----	118

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

I, Dr. Michael T. Johnson, hereby declare:

I. INTRODUCTION

1. I have been retained by Meta Platforms, Inc. (“Petitioner”) to provide my opinions concerning the validity of claim 1-46 of U.S. Patent No. 8,503,691 (“the ’691 patent”).

2. I have been simultaneously engaged by Petitioner to provide a declaration for IPR petition challenging U.S. Patent No. 10,779,080.

3. This Declaration summarizes the opinions I have formed to date. I reserve the right to modify my opinions, if necessary, based on further review and analysis of information that I receive subsequent to the filing of this report, including in response to positions taken by Jawbone Innovations LLC (“Jawbone”) or its experts that I have not yet seen, including any secondary consideration evidence that Jawbone or its expert may consider and present.

4. It is my opinion that claims 1-46 of the ’691 patent are invalid based on the following grounds:

Ground #	Claims Challenged	Ground
1	1-46	Obvious under §103 from Brandstein
2	12-14, 17-19, 28, 38, 39	Obvious under §103 from Brandstein and McCowan

II. EXPERIENCE AND QUALIFICATIONS

5. I am currently the Associate Dean for Undergraduate Education and Student Success of the University of Kentucky (UK) Pigman College of Engineering in Lexington, Kentucky, and have served in that position since July 2023. I also currently serve as Department Chair of Electrical and Computer Engineering (ECE) at UK and have been a Full Professor in ECE since 2016. Prior to that, I was a faculty member at Marquette University in Milwaukee, Wisconsin from 2000 to 2016, during which time I also spent two separate sabbatical years as a Visiting Professor at Tsinghua University in Beijing, China.

6. In 2000, I received my PhD from Purdue University in West Lafayette, Indiana with a doctoral dissertation focused on speech processing and automatic speech recognition. My undergraduate degrees are in Computer Science Engineering and in Electrical Engineering from LeTourneau University in Longview, TX, in 1989 and 1990, respectively, and I received a Master's degree in Electrical Engineering from the University of Texas at San Antonio in 1994. Prior to receiving my PhD, I worked in industry as a design engineer and engineering manager for more than 5 years, primarily in the telecommunications industry.

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

7. The focus of my research work is Speech and Audio Signal Processing, and I regularly conduct and publish original research articles within this field, with a publication record including 50 peer-reviewed journal articles and more than 140 total peer-reviewed publications over the past two decades. My specific areas of expertise include digital signal processing, speech enhancement and noise reduction, automatic speech recognition, speech production and kinematics, bioacoustics, and machine learning. I have received multiple research grant awards for my work in these areas, including from the National Science Foundation and the National Institutes of Health.

8. As a faculty member, I regularly teach both undergraduate and graduate courses in electrical and computer engineering. This includes undergraduate level courses, such as Circuits, Signals and Systems, and Digital Signal Processing, as well as graduate level courses, such as Optimal and Adaptive Signal Processing and Speech Processing.

9. My professional service includes active membership in the Institute for Electrical and Electronic Engineers (IEEE), the IEEE Signal Processing Society, the Acoustical Society of America (ASA), and the Electrical and Computer Engineering Department Heads Association (ECEDHA). I served as a member of the IEEE Signal Processing Society Speech and Language Technical Committee from 2014 to

2017, and regularly serve as a reviewer for journals, conferences, and workshops in the area of speech and signal processing, including IEEE Transactions on Acoustics, Speech, and Signal Processing, IEEE Transactions on Signal Processing, the Journal of the Acoustical Society of America, Speech Communication, the International Conference on Acoustics, Speech, and Signal Processing, the IEEE Workshop on Automatic Speech Recognition, and Interspeech. I have reviewed funding proposals for a number of public and private agencies and foundations, including for multiple programs and directorates at the National Science Foundation. In addition to these manuscript and grant proposal reviewing activities, I also currently serve as Area Editor for the journal Speech Communication in the area of speech enhancement.

10. During my academic career, I have been honored to receive a number of awards for my teaching and research work. Notable awards have included the College of Engineering Researcher of the Year at Marquette University in 2007, Milwaukee Young Engineer of the Year in 2006, and Electrical and Computer Engineering Eta Kappa Nu Teacher of the Year in 2005 and 2014. In 2008, my work applying speech recognition technology to animal vocalizations was featured in the Milwaukee Journal-Sentinel series “Groundbreaking Thinkers in Wisconsin.”

11. A detailed record of my professional qualifications, including a list of publications, awards, and professional activities, is set forth in my curriculum vitae, which is attached as Exhibit 1004.

III. UNDERSTANDING OF RELEVANT LEGAL PRINCIPLES

A. Claim Construction

12. I understand that claim construction begins with the language of the claims. I understand that words are given their “ordinary and customary meaning,” which is “the meaning that the term would have to a person of ordinary skill in the art in question at the time of the invention, i.e., as of the effective filing date of the patent application.” I further understand that specification and prosecution history are sources of intrinsic evidence that informs the meaning of the claim language. Extrinsic evidence, such as expert testimony and dictionaries, may be useful in educating the court regarding the field of the invention or helping determine what a person of ordinary skill in the art (“POSITA”) would understand claim terms to mean. I understand that extrinsic evidence in general is viewed as less reliable than intrinsic evidence.

13. I understand that in this proceeding, any claim term that is not construed should be given its plain and ordinary meaning. For purposes of this IPR only, I

have been asked to assume that the '691 patent claim terms do not require express construction and can be afforded their plain and ordinary meaning.

14. Additionally, I understand that during claim construction in *Jawbone Innovations, LLC v. Google LLC*, C.A. No. 21-cv-00985 (W.D. Tex. Sept. 23, 2021), the Court determined that the terms “substantially similar” and “substantially dissimilar” are indefinite. I have been asked, for the purposes of this declaration only, to assume the claims are not invalid under § 112. For the purposes of this declaration only, I interpret “substantially similar” responses to include nearly identical, or constant, responses, and “substantially dissimilar” responses to include very different responses.

B. Obviousness

15. I understand that a prior art reference can render a patent claim obvious to one of ordinary skill in the art if the differences between the subject matter set forth in the patent claim and the prior art are such that the subject matter of the claim would have been obvious at the time the claimed invention was made. In analyzing obviousness, I understand that it is important to consider the scope of the claims, the level of skill in the relevant art, the scope and content of the prior art, the differences between the prior art and the claims, and any secondary considerations.

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

16. I understand that when the claimed subject matter involves combining pre-existing elements to yield no more than one would expect from such an arrangement, the combination is obvious. I also understand that in assessing whether a claim is obvious one must consider whether the claimed improvement is more than the predictable use of prior art elements according to their established functions. I understand that there need not be a precise teaching in the prior art directed to the specific subject matter of a claim because one can take account of the inferences and creative steps that a person of skill in the art would employ. I further understand that a person of ordinary skill is a person of ordinary creativity, not an automation.

17. I understand that obviousness cannot be based on the hindsight combination of components selectively culled from the prior art. I understand that in an obviousness analysis, neither the motivation nor the avowed purpose of the inventors controls the inquiry. Any need or problem known in the field at the time of the invention and addressed by the patent can provide a reason for combining elements. For example, I understand that it is important to consider whether there existed at the time of the invention a known problem for which there was an obvious solution encompassed by the patent's claims. I understand that known techniques can have obvious uses beyond their primary purposes, and that in many cases a person of

ordinary skill can fit the teachings of multiple pieces of prior art together like pieces of a puzzle.

18. I understand that, when there is a reason to solve a problem and there are a finite number of identified, predictable solutions, a person of ordinary skill has good reason to pursue the known options within his or her technical grasp. I further understand that, if this leads to the anticipated success, it is likely the product not of innovation but of ordinary skill and common sense, which bears on whether the claim would have been obvious.

19. I understand that secondary considerations can include, for example, evidence of commercial success of the invention, evidence of a long-felt need that was solved by an invention, evidence that others copied an invention, or evidence that an invention achieved a surprising result. I further understand that such evidence must have a nexus, or causal relationship to the elements of a claim, in an order to be relevant. I am unaware of any such secondary considerations for the '691 patent.

IV. MATERIALS REVIEWED AND CONSIDERED

20. All of the opinions contained in this Declaration are based on the documents I reviewed and my knowledge and professional judgment. In forming the

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

opinions expressed in this Declaration, while drawing on my experience in the field of speech and audio signal processing, I reviewed the following documents:

Exhibit No.	Description
1001	U.S. Patent No. 8,503,691 to Burnett (“the ’691 patent”)
1002	File History of U.S. Patent No. 8,503,691
1005	Excerpts of MICROPHONE ARRAYS: SIGNAL PROCESSING TECHNIQUES AND APPLICATIONS (Michael Brandstein & Darren Ward eds., Springer-Verlag 2001) (“Brandstein”)
1006	Lloyd Griffiths & Charles Jim, <i>An Alternative Approach to Linearly Constrained Adaptive Beamforming</i> , vol. AP-30, no. 1 IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, 27 (Jan. 1982) (“Griffiths-Jim”)
1009	Iain A. McCowan et al., <i>Near-Field Adaptive Beamformer for Robust Speech Recognition</i> , vol. 12, no. 1 DIGITAL SIGNAL PROCESSING, 87 (Jan. 2002) (“McCowan”)
1010	U.S. Patent No. 5,473,701 (“Cezanne”)
1011	U.S. Patent No. 5,651,071 (“Lindemann”)
1012	U.S. Patent No. 5,627,799 (“Hoshuyama”)
1013	U.S. Patent No. 7,386,135 (“Fan”)
1014	Claim Construction Order, ECF No. 88, <i>Jawbone Innovations, LLC v. Google LLC</i> , C.A. No. 21-cv-00985 (W.D. Tex.)

V. LEVEL OF SKILL IN THE ART AND PERSPECTIVE APPLIED IN THIS DECLARATION

21. I understand that certain issues relating to validity must be judged from the perspective of a person of ordinary skill in the relevant art, as I discuss below.

22. The invention of the '691 patent generally relates to noise suppression systems, devices, and methods for use in acoustic applications.

23. In determining the level of ordinary skill, I have been asked to consider, for example, the types of problems encountered in the field, prior solutions to those problems, the rapidity with which innovations are made, the sophistication of the technology, and the educational level of active workers in the field. Taking those factors into consideration, a person of ordinary skill in the art at the time of the claimed priority date of the '691 patent would have had a minimum of a bachelor's degree in computer engineering, computer science, electrical engineering, mechanical engineering, or a similar field, and approximately three years of industry or academic experience in a field related to acoustics, speech recognition, speech detection, or signal processing. In my opinion, work experience could substitute for formal education, and additional formal education could substitute for work experience.

24. My opinions do not turn on the precise definition, and the challenged claim would be unpatentable from the perspective of any reasonable POSITA.

25. I believe that I would have qualified as a person of at least ordinary skill in the art as of the earliest claimed June 13, 2007 priority date on the face of the '691 patent. I believe that I have a sufficient level of knowledge, experience, and education to provide an expert opinion in the field of the '691 patent.

26. My opinions in this Declaration are based on the perspective of a person of ordinary skill in the art as of June 13, 2007. This is true even if the testimony is stated in the present tense. Each of the statements below reflects my opinion based on my review of the disclosures of the '691 patent, the file history of the '691 patent, and the challenged claims.

VI. BACKGROUND

27. As of the earliest possible priority date of the '691 patent in 2007, techniques for reducing noise in signals using multiple microphones had been known for decades. One prominent technique was the Generalized Sidelobe Canceler, or GSC. The GSC is sometimes also called the Griffiths-Jim beamformer ("GJBF") after its inventors.

A. In 1982, Griffiths and Jim Publish Their Seminal Article on the GSC.

28. In 1982, Lloyd Griffiths and Charles Jim published an article in IEEE Transactions on Antennas and Propagation describing “a simple time-varying beamformer which can be used to combine the outputs of an array of sensors.” Ex. 1006, 27. The purpose of the beamformer was “to minimize the effects of noise and interference at the array output” while capturing the target signal. Ex. 1006, 27.

29. Griffiths and Jim called their beamformer a “generalized sidelobe canceling” structure. *Id.*, 29. As shown in Fig. 4 of their paper, the signal processor had two main substructures: a top branch that was a “conventional beamformer” designed to capture the target signal plus noise, and a bottom branch that was the “sidelobe canceling path” designed to capture only noise so that the noise could be subtracted or canceled out:

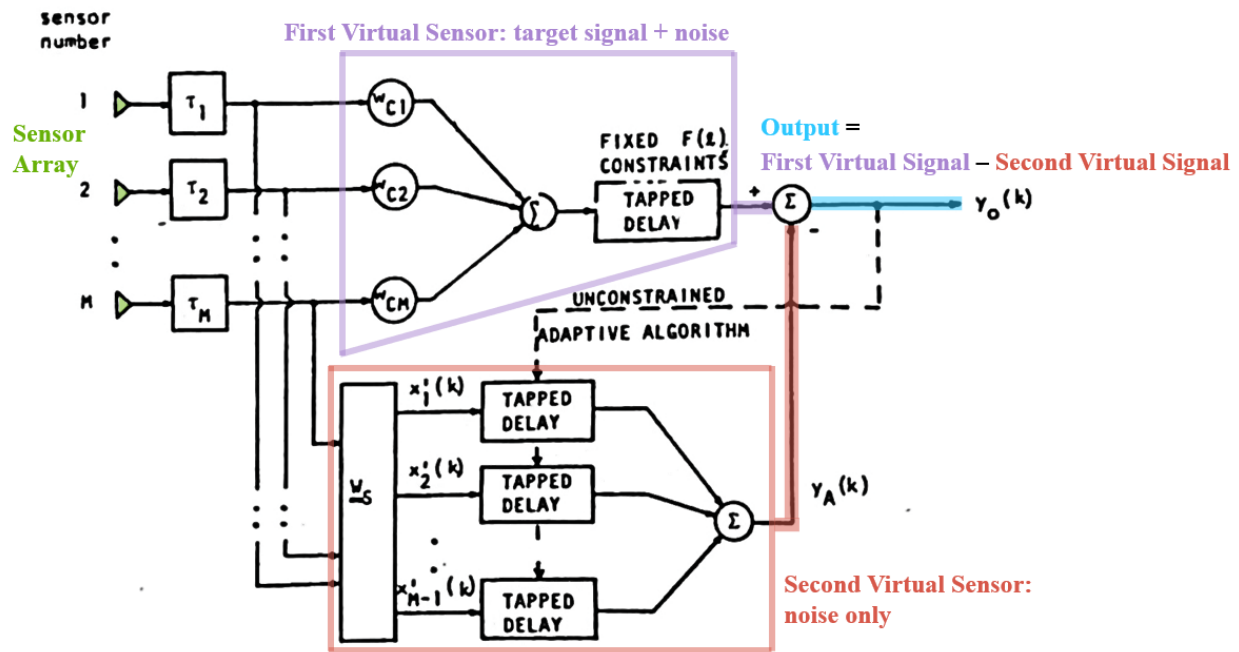


Fig. 4. Generalized sidelobe canceling form of linearly constrained adaptive array processing algorithm.

Ex. 1006, Fig. 4 (annotated)

Ex. 1006, 29–30.¹ In the top branch of the GSC, the outputs of the **array sensors** are combined to form a conventional beamformer, which I will refer to as the **first virtual sensor**. The sensor outputs are delayed by “time-delay steering elements $\tau_1, \tau_2, \dots, \tau_M$ ” and then are combined by multiplying the sensor output signals by factors called “weights” ($w_{c1}, \dots, w_{cM}$ in the paper). These “weights” are also sometimes also called “gains”. The weighted sensor signals are further summed. Ex. 1006, 29–30.

¹ I have annotated the figures with color throughout.

Griffiths and Jim denote the output of the first virtual sensor as $y_c'(k)$ in the paper's equations. Ex. 1006, 30. This output contains the target signal and noise.

30. As Griffiths-Jim explains, “[t]he lower path in Fig. 4 is the sidelobe canceling path.” Ex. 1006, 30. I refer to this as the **second virtual sensor**. Like the first virtual sensor, the second virtual sensor is formed from a combination of the outputs of the array sensors. However, the combination used for the second virtual sensor differs from the combination used for the first virtual sensor. The lower path includes a blocking matrix (BM) designed to “block the desired signal $s(k)$ from the lower path.” Ex. 1006, 30. The array sensor outputs are combined by filtering and summing (including by blocking the desired signal), and the output of the second virtual sensor is denoted in the paper as $y_A(k)$. Ex. 1006, 30. The second virtual sensor's output “ $y_A(k)$ contains no desired signal terms” and instead “contains only noise and interference terms.” Ex. 1006, 30.

31. By subtracting the noise-only output of the second virtual sensor from the target-and-noise output of the first virtual sensor, the overall **output** of the GSC, $y(k)$, is obtained:

$$y(k) = y_c'(k) - y_A(k). \quad (29)$$

Ex. 1006, 30. The resulting cleaned-up signal contains reduced noise without distorting the desired signal. Ex. 1006, 30 (output due to desired signal satisfies the constraint defined by paper's equation 9, which defines a constraint for "zero distortion" (p. 28)).

B. The GSC Was Applied to Microphone Arrays to Reduce Noise

32. In the two decades following Griffiths and Jim's article, the GSC was applied to many signal-processing applications, including microphone arrays to reduce noise in speech applications. For example, U.S. Patent No. 5,651,071 ("Lindemann"), filed in 1993, cites the article and explains that it was known that a GJBF could be used "to improve signal-to-noise ratio for hearing aids" was known. Ex. 1011, 1:40-46, 12:12-14.² As another example, Griffiths-Jim is the first non-patent reference cited in U.S. Patent No. 5,627,799 ("Hoshuyama '799 patent"), filed in 1995, which relates to "interference cancelers, and more particularly to a generalized sidelobe canceler, or adaptive beamformer for an array of sensors such as micro-

² Patent citations are in column:line format.

phones[.]” Ex. 1012, 1:8-11. Describing what was prior art even at that time, Hoshuyama explains that one way the GSC had been used with microphone arrays was as follows:

According to a prior art microphone array, signals detected by an array of microphones are lowpass filtered and summed together to detect a target signal that arrives in a particular direction. The adaptive microphone array beamformer is one form of the generalized sidelobe canceler as described in an article “An alternative Approach to Linearly Constrained Adaptive Beamforming”, Lloyd J. Griffiths and Charles W. Jim, the IEEE Transactions on Antenna and Propagation, Vol. AP-30, No. 1, January 1982, pages 27-34.

Ex.1012, 1:17-26.

C. A 2001 Book Edited by Brandstein Illustrates How to Apply a GSC to a Microphone Arrays to Reduce Noise

33. Published in 2001, MICROPHONE ARRAYS: SIGNAL PROCESSING TECHNIQUES AND APPLICATIONS (Ex. 1005 (“Brandstein”)), provides “a single complete reference on microphone arrays.” Ex. 1005, Preface. The book became a standard reference for those in the field of audio-signal processing.

34. At the outset of Chapter 5, which discusses robust adaptive beamforming, Brandstein explains that “[a]pplications of beamforming include microphone arrays for speech enhancement.” Ex. 1005, 87 (original page numbering). As Brandstein explains, “[t]he goal of speech enhancement is to remove undesirable signals such as noise and reverberation.” Ex. 1005, 87. Brandstein further explains

that, among various known adaptive beamformers, “the Griffiths-Jim beamformer (GJBF), or the generalized sidelobe canceler, is most widely known.” Ex. 1005, 88 (internal citation omitted). In Fig. 5.1, Brandstein depicts the structure of the basic GJBF. Ex. 1005, 88.

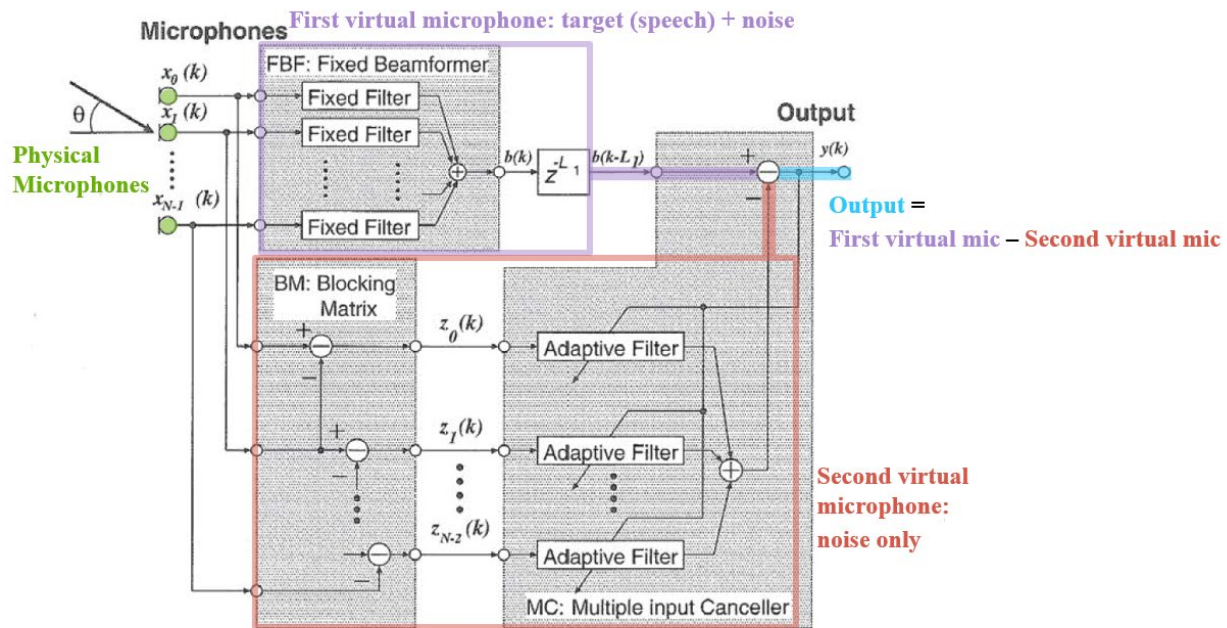


Fig. 5.1. Griffiths-Jim beamformer. It comprises a fixed beamformer (FBF), a multiple-input canceler (MC), and a blocking matrix (BM).

Ex. 1005, Fig. 5.1 (annotated)

As shown, the signals of a first and second physical microphone, $x_0(k)$ and $x_1(k)$, are combined in the top branch—a first virtual microphone—by filtering and summing the signals to form a fixed beamformer (FBF). The first virtual microphone captures the target speech signal plus noise.

35. In the bottom branch, the signals of the two **physical microphones** are combined by summing and filtering the signals in a different way to form a **second virtual microphone**, which is depicted in the annotated Fig. 5.1 above including the blocking matrix. As Brandstein explains, “the BM forms a null in the look direction so that the target signal is suppressed and all other signals are passed through.” Ex. 1005, 88. Brandstein further explains that “[t]he BM was named after its function, which is to block the target signal.” Ex. 1005, 88. As a result, the second virtual microphone captures noise only.

36. The overall **output** is the speech-and-noise output of the first virtual microphone minus the noise-only output of the MC. As a result, “in the subtracter output $y(k)$, the target signal is enhanced and undesirable signals such as ambient noise and interferences are suppressed.” Ex. 1005, 88-89; *see also* 33-34 (explaining that the lower branch of the GJBF is “subtracted from the output of the upper part of the structure to get the noise-reduced output signal.”).

37. The two virtual microphones have very different responses to the target signal (i.e., speech). The first virtual microphone is designed to capture the target speech signal. On the other hand, the second virtual microphone is designed to block

it. In contrast, they have nearly the same responses to noise so that the final subtraction removes the noise and enhances the target speech. Brandstein illustrates this in Fig. 5.2, which shows the directivity pattern for the final output of an example GJBF:

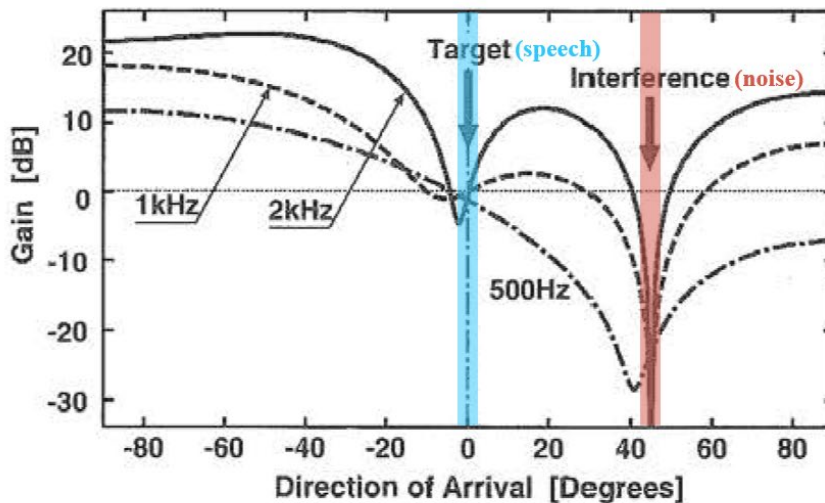


Fig. 5.2. Example directivity pattern of the Griffiths-Jim beamformer.

Ex. 1005, Fig. 5.2 (annotated)

The horizontal axis of the graph shows direction of arrival (“DOA”) measured in degrees relative to the direction perpendicular to the axis of the microphone array: the **target (speech) signal** is at 0 degrees, while the **interference (noise) signal** is at approximately 45 degrees. The vertical axis of the graph shows gain in decibels: zero gain corresponds to no change in sound pressure or signal power, while a positive gain corresponds to an increase in signal power, and a negative gain corresponds to a decrease in signal power.

38. As highlighted in blue, the target (speech) signal is reproduced faithfully with essentially zero gain, reflecting that subtracting the second virtual microphone's noise-only output from the first virtual microphone's speech-and-noise output will produce an unchanged target (speech) signal in the direction of the target (speech). On the other hand, as highlighted in red, Fig. 5.2 shows a highly negative gain in the direction of the interference (noise) signal, reflecting that subtracting the two virtual microphones' outputs cancels the noise. Ex. 1005, 90. As Brandstein explains, "[i]n the direction of the target signal, almost constant gains close to 0 dB are obtained over a wide range of frequencies. On the contrary, in the direction of the interference, a deep null is formed."

39. Brandstein further provides the individual directivity pattern for the FBF and BM components of the GJBF in Fig. 5.3. These directivity patterns show non-zero responses to the interference (noise) in the direction of the interference. The directivity patterns also show that BM response has a null in the direction of the target (speech) while the FBF response has a non-zero response to the target in that direction.

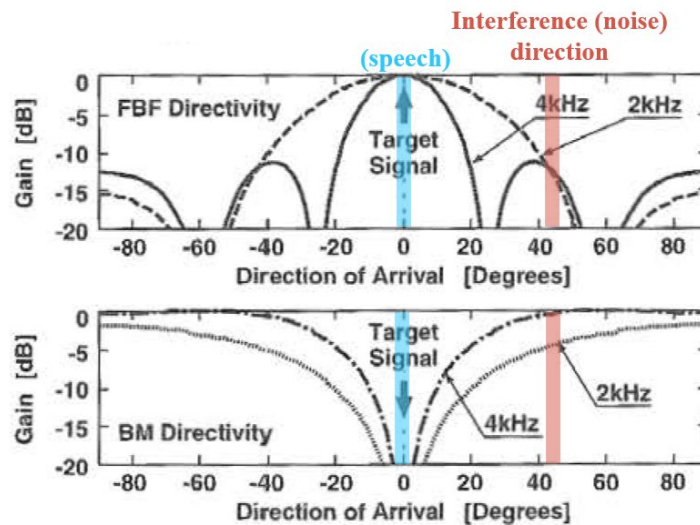


Fig. 5.3. Directivity pattern of a fixed beamformer (FBF) and a blocking matrix (BM).

Ex. 1005, Fig. 5.3 (annotated)

40. According to Brandstein, the basic GJBF depicted in Fig. 5.1 may have errors in the steering-vector, leading to undesirable “leakage” of the target in the BM and potential cancellation of the target signal at the output. Ex. 1005, 90-91, 97. To reduce this undesirable effect, Brandstein teaches ways for the GJBF to (1) track the direction of the target signal, and (2) adapt the BM coefficients only when the target signal is absent, thereby minimizing the effect of target leakage on performance of the beamformer. Ex. 1005, 90-91, 97. To accomplish this, Brandstein incorporates adaptive filters into the bottom branch of the GJBF (containing the *second virtual*

microphone), coupled with an adaptation-mode controller (“AMC”) to detect presence or absence of the target signal. Ex. 1005, 90-91, 97. Brandstein refers to this implementation as the “GJBF-CNA,” which is depicted in Fig. 5.8:

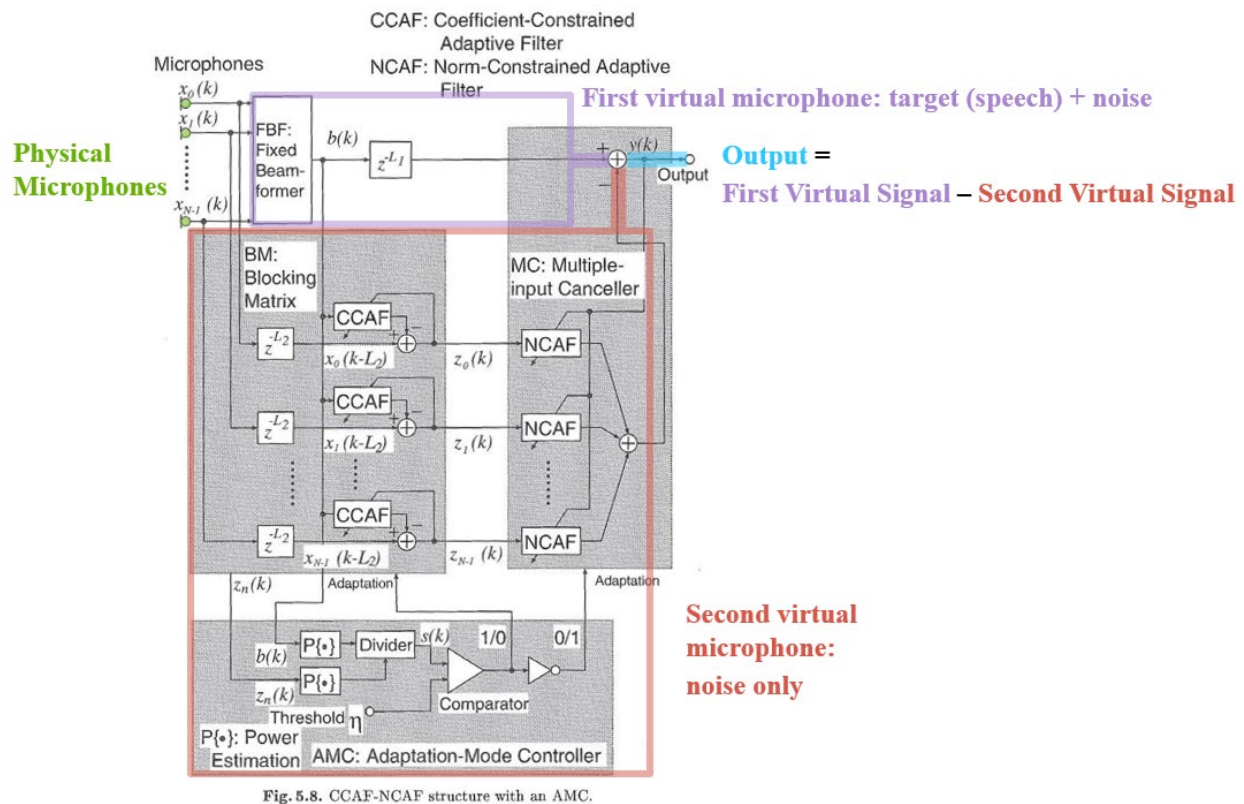


Fig. 5.8. CCAF-NCAF structure with an AMC.

Ex. 1005, Fig. 5.8 (annotated)

41. Brandstein depicts the “sensitivities to the frequency component of the target signal” for the GJBF-CNA, which are plotted in Fig. 5.10 (below). Fig. 5.10 shows how the sensitivities depend on different frequencies (500, 1000, 2000, 3000, and 3500 Hz). Ex. 1005, 100. Brandstein explains that the GJBF-CNA is suitable

for broadband applications such as microphone arrays because the frequency dependency is small throughout a substantial angular region in the direction of the target source, as shown in Fig. 5.10. Ex. 1005, 100.

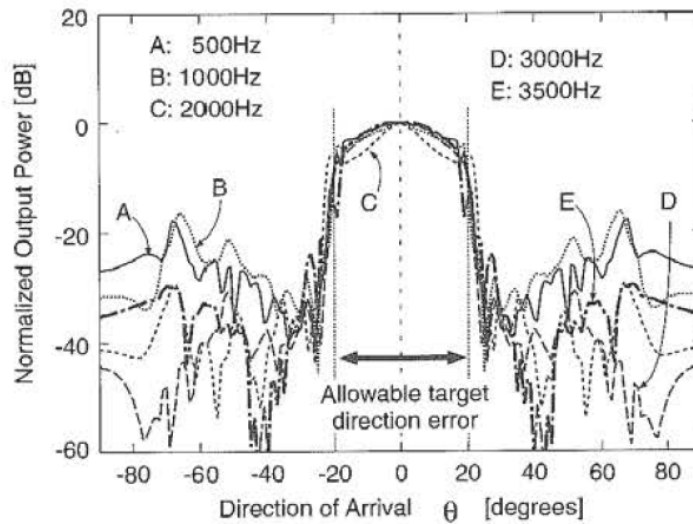


Fig. 5.10. Sensitivities after convergence as a function of DOA at different frequencies.

Ex. 1005, Fig. 5.10

D. Near-field Compensation Techniques Were Well Known

42. Delaying signals based on geometry to adjust for differences in arrival times was well known. Adaptive beamforming relies on the spatial geometry of the array and the signal's direction of arrival. Ex. 1005, 87. When the source of the target signal is close to the array, instead of treating the signal's wavefront as a plane, it may be useful to account for the spherical geometry of the wavefront—i.e., to use

a *near-field* design. Near-field compensation techniques were well known as of June 13, 2007 and used basic calculations to determine the delays between signal arrival at the various microphones in the array. Ex. 1005, 87. Brandstein discloses the standard near-field design:

- Nearfield design, including attenuation of the desired signal [14], [22]

$$\mathbf{d}^T = [a_0 \exp(-j\omega\tau_0), a_1 \exp(-j\omega\tau_1), \dots, a_{N-1} \exp(-j\omega\tau_{N-1})] , \quad (2.28)$$

$$a_i = \frac{\|\mathbf{q} - \mathbf{p}_{ref}\|}{\|\mathbf{q} - \mathbf{p}_i\|} , \quad (2.29)$$

$$\tau_i = \frac{\|\mathbf{q} - \mathbf{p}_{ref}\| - \|\mathbf{q} - \mathbf{p}_i\|}{c} , \quad (2.30)$$

where $\|\mathbf{q} - \mathbf{p}_{ref}\|$ and $\|\mathbf{q} - \mathbf{p}_i\|$ denote the distance between the vector location of the source $\mathbf{q}$ and a reference sensor $\mathbf{p}_{ref}$, or the sensor $\mathbf{p}_i$, respectively.

Ex. 1005, 25. The time it takes for a signal to travel from a source to a microphone is the distance traveled divided by the speed of sound, c . The τ_i terms above represent the time difference for a signal to propagate from the source to microphone i compared to a reference microphone. (*Id.*) The complex exponential shown above includes the τ_i time difference terms. By applying that exponential to the signal received at the reference microphone, the received signal is delayed to reproduce (in terms of phase) the signal as received at microphone i . Applying the a_i term to the reference signal adjusts the amplitude of the signal to match the amplitude of the

signal as it is received at microphone i . The purpose of applying the a_i term is to ensure that the amplitudes are the same across the different channels. The signal x_i on microphone i can be expressed in terms of the signal x_{ref} on the reference microphone as follows (where $j = \sqrt{-1}$):

$$x_i(\omega) = a_i(e^{j\omega\tau_i})x_{ref}(\omega).$$

43. The math relevant to near-field compensation is further explained in the January 2002 edition of the journal DIGITAL SIGNAL PROCESSING, in an article by McCowan et al. titled *Near-Field Adaptive Beamformer for Robust Speech Recognition* (“McCowan”). Ex. 1007, 89-90. As I explain below, McCowan expressly motivates and teaches incorporating the near-field design into the GSC. *Id.*, 90-96.

VII. THE '691 PATENT

A. Specification

44. The '691 patent purports to distinguish itself from prior art “conventional arrays and algorithms” that “seek to reduce noise by nulling out noise sources” by using a virtual microphone with a null in the direction of speech “to remove the speech of the user.” Ex. 1001, 4:61-5:2. However, in my opinion, this is exactly what a GSC does. Specifically, a GSC uses a virtual microphone with a null in the direction of the target signal to obtain a noise-only reference.

45. The '691 patent describes a “dual omnidirectional microphone array (DOMA)” for adaptive noise suppression. Ex. 1001, 4:61-62, 5:63-64. Two omnidirectional microphones are used to “form two distinct virtual directional microphones which are configured to have very similar noise responses and very dissimilar speech responses.” Ex. 1001, Abstract, 4:61-67, 5:59-62, 8:46-58. The two virtual microphones are “paired with an adaptive filter algorithm” to reduce noise. Ex. 1001, 5:2-6.

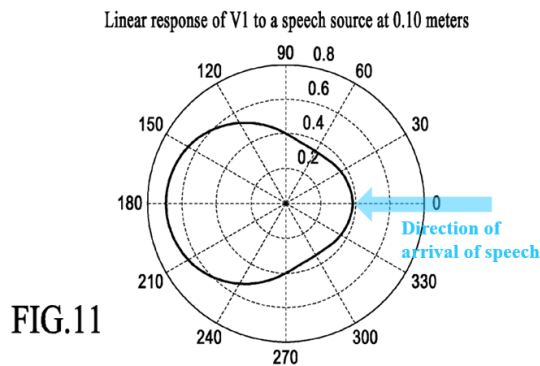
46. In one embodiment, two virtual directional microphones V1 and V2 are formed based on two physical omnidirectional microphone outputs O1 and O2 according to certain parameters expressed in the '691 patent as:

$$\begin{aligned}\beta &= \frac{d_1}{d_2} \\ \gamma_- &= \frac{d_2 - d_1}{c} \cdot f_x \text{ (samples)} \\ d_1 &= \sqrt{d_S^2 - 2d_S d_0 \cos(\theta) + d_0^2} \\ d_2 &= \sqrt{d_S^2 + 2d_S d_0 \cos(\theta) + d_0^2}\end{aligned}$$

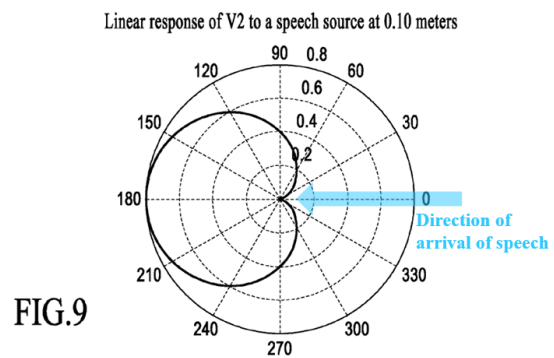
Ex. 1001, 10:59-66. The '691 patent explains that “[t]he distances d_1 and d_2 are the distance from O_1 and O_2 to the speech source (see FIG. 2), respectively, and γ is their difference divided by c , the speed of sound, and multiplied by the sampling

frequency f_s .” Ex. 1001, 11:1-6. Thus, γ is simply the difference in propagation time to the microphones taken from the standard near-field design equations (denoted τ_i in ¶¶29, 42, above), just expressed in number of samples instead of other units of time, like seconds. Similarly, β is simply the amplitude-adjustment factor from the standard near-field design (denoted a_i in ¶42, above).

47. According to the '691 patent, using these metrics for V_1 and V_2 produces desirable directivity patterns with respect to a speech signal expected to arrive at an angle of zero degrees from 10 centimeters away, as illustrated in Figs. 9 and 11:



Ex. 1001, Fig. 11



Ex. 1001, Fig. 9

As shown, V_1 (Fig. 11) exhibits a response to the speech signal, while V_2 (Fig. 9) has a null in the expected direction of speech “and will therefore exhibit minimal response to the speech.” Ex. 1001, 11:30-32, 12:20-24. In contrast, the '691 patent

explains that V_1 and V_2 will supposedly have “very similar” responses to noise originating a meter away from the array. Ex. 1001, 12:25-29.

48. The '691 patent describes the linear response of virtual microphone V_1 as “flat for all frequencies within approximately ± 30 degrees of the axis of the array,” as shown in Fig. 13 (reproduced below). Ex. 1001, 12:30-35, 12:43-61.

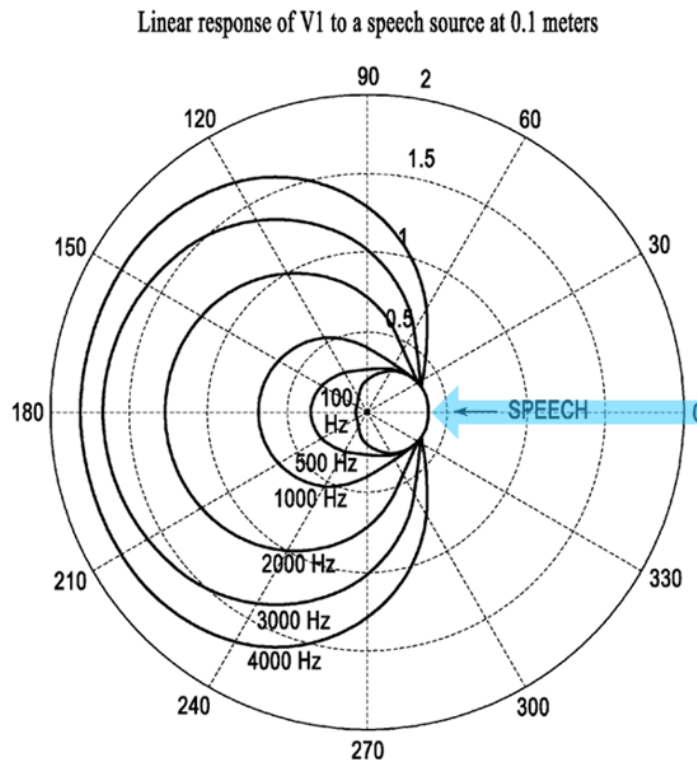


FIG.13

Ex. 1001, Fig. 13 (annotated).

B. Prosecution History

49. I understand that the '691 patent was filed as U.S. Patent Application No. 12/139,333 (“the '333 Application”) and claims priority to four provisional applications, the earliest of which was filed June 13, 2007. During prosecution of the '333 Application, the Examiner rejected pending claims 1, 2, 21, 22, and 30 as anticipated by U.S. Patent No. 5,473,701 (Ex. 1010, “Cezanne”) and the remaining pending claims as obvious over Cezanne and U.S. Patent No. 7,386,135 (“Fan”). Ex. 1002, 133-135.

50. I understand that in response, Applicant amended pending claim 1 to add “the first virtual microphone having a first linear response to speech, the first linear response being substantially similar across a plurality of frequencies for a speech source located within a predetermined angle relative to an axis of the microphone array,” and made similar amendments to the other pending independent claims. Ex. 1002, 178, 181-184, 186. Applicant argued that Cezanne and Fan did not disclose or render obvious this limitation. Ex. 1002, 166-167.

51. The Examiner rejected the pending claims based on the same grounds. Ex. 1002, 191-193. The Examiner reasoned that Cezanne disclosed a first virtual microphone that disclosed the recited first linear response to speech because

Cezanne disclosed equations that were similar to those disclosed in the specification of the '333 Application. Ex. 1002, 193; *see* Ex. 1010, 5:16-33; Ex. 1002, 16, 20.

52. In response, Applicant amended pending claim 1 to add (1) “a first linear response to noise,” (2) “the second virtual microphone having a second linear response to speech and a second linear response to noise, the second linear response to noise being substantially similar to the first linear response to noise, and the second linear response to speech being substantially dissimilar to the first linear response to speech,” and (3) “at least one of the first linear response to noise and the second linear response to [noise] is being non-zero in a direction toward a source of noise,” and amended claim 23 to require that the “first linear response to noise [] is devoid of a null.” Ex. 1002 at 204, 208. Applicant made similar amendments to the other pending independent claims. Ex. 1002, 208-212, 214-215. To distinguish the prior art, Applicant emphasized what the prior art disclosures lacked, arguing that Cezanne and Fan did not disclose or suggest a first or second linear response to noise that is “non-zero in a direction toward a source of noise.” Ex. 1002, 217-220.

53. The Examiner then issued a notice of allowance, stating that “a dual virtual microphone array as claimed including having a linear response to both speech and noise could not be anticipated nor, in combination, be rendered obvious over the prior art of record.” Ex. 1002, 235-236.

C. The Earliest Possible Priority Date of the '691 Patent Claims Is June 13, 2007.

54. The earliest application to which the '691 patent claims priority is U.S. provisional application 60/934,551. I understand that the earliest possible priority date for all challenged claims is the '551 application's filing date, June 13, 2007.

D. The References Are Prior Art.

55. I understand that all references relied on in the grounds (§ 4, above) are prior art under pre-AIA §102(b) because they published more than one year before June 13, 2007:

- (1) Brandstein published no later than 2001 (Ex. 1005); and
- (2) McCowan published no later than 2002 (Ex. 1007).

E. The Asserted References Are Analogous Art.

56. I understand that prior art is analogous if (1) it is from the same field of endeavor, regardless of the problem addressed, or (2) it is reasonably pertinent to the particular problem with which the inventor is involved. The references I rely on in the grounds are analogous art because, in my opinion, they are from the same field as the '691 patent, e.g., microphone array signal processing and noise reduction. Ex.

1001, 1:14-17, 4:61-62. Additionally, in my opinion, those references are also pertinent to a problem the inventor was focused on, e.g., noise suppression. Ex. 1001, 1:14-17, 4:61-62.

VIII. GROUNDS OF UNPATENTABILITY

57. For the reasons I explain in greater detail below, it is my opinion that as of the priority date of the '691 patent on its face (June 13, 2007), a POSITA would have considered the claims of the '691 patent unpatentable as obvious over the prior art discussed in this Declaration and listed in the Grounds above in ¶4.

A. Ground 1: Brandstein Renders Obvious Claims 1-46

58. For the reasons I explain below, it is my opinion that claims 1-46 would have been obvious in view of Brandstein.

1. Motivation to Combine

59. Brandstein teaches two embodiments of the GJBF: a basic GJBF (depicted in Figs. 5.1-5.3) and a proposed GJBF (or GJBF-CNA, depicted in Figs. 5.8-5.10). While both embodiments are similar (with identical top branches including the same FBF), the bottom branch of the proposed GJBF is more sophisticated, in that it includes delay modules and adaptive filters in the BM, norm-constrained adaptive filters in the MC, and an adaptation-mode controller (AMC) to control adaptation of the BM and MC. (*Compare* Fig. 5.1 with Fig. 5.8; *see also* Ex. 1005, 88-

90, 97-99.) While Brandstein does not expressly disclose that the simple GJBF's directivity pattern for the top branch that includes the FBF (*the first linear response*) is *substantially similar across a plurality of frequencies*, in my opinion, this would have been obvious to the POSITA based on the teachings of Brandstein related to the proposed GJBF and the knowledge of the POSITA.

60. Specifically, Brandstein's GJBF-CNA (Fig. 5.8) meets this limitation because the corresponding response depicted in Fig. 5.10 depicts a "constant" sensitivity for different frequencies. Ex. 1005, 99-100, Fig. 5.10. A POSITA would have been motivated to modify Brandstein's basic GJBF with Brandstein's GJBF-CNA structure, so that the *first linear response to speech* is constant *across a plurality of frequencies*.

61. It is my opinion that a POSITA would have been motivated to apply Brandstein's teachings related to the GJBF-CNA to the simple GJBF. Brandstein teaches that target-signal cancellation due to steering-vector errors may occur in the basic GJBF, which potentially leads to undesirable "leakage" of the target in the BM and target signal cancellation at the output. Ex. 1005, 90-91, 97. To reduce this undesirable effect, Brandstein incorporates adaptive filters into the bottom branch of the GJBF-CNA, coupled with an AMC to detect presence or absence of the target signal. *Id.*, 87-90, Fig. 5.8. As a result, the modified GJBF can (1) track the direction

of the target signal, and (2) adapt the BM coefficients only when the target signal is absent, so that any target leakage has a minimal effect on performance of the beamformer. Ex. 1005, 90-91, 97. In my opinion, a POSITA would have been motivated to apply these modifications, to reduce target cancellation.

62. Moreover, because both of Brandstein's embodiments share the same goals and have the same overall beamforming structure (*e.g.*, subtract a noise-only signal from a target plus noise signal using spatial channels), the POSITA would have understood that the overall shapes and characteristics of the response patterns of the basic GJBF's top and bottom branches (including the FBF and BM) are unchanged when the basic GJBF is modified.

63. For example, comparing Figs. 5.1 and 5.8, the POSITA would have appreciated that the FBF is the same between the simple and proposed GJBFs (Ex. 1005, 89, 98), such that the response of the top branch (containing both speech and noise) is the same in both.

64. The POSITA would have known that the entire purpose of modifications to the bottom branch of the GJBF is to make it more robust to steering vector errors toward the direction of the target without changing the response of the BM elsewhere. As Brandstein depicts in Figs. 5.9 and 5.10, the proposed GJBF also has the benefit of having a wider region that is responsive to the target *speech* signal,

thus achieving a “robustness against 20° target-direction error and high interference-reduction performance.” Ex. 1005, 100. The POSITA would have understood that this wider region is achieved due to the proposed GJBF’s BM, which incorporates adaptive filters to “provide[] a wider null for the target with sharper edges.” Ex. 1005, 97. A POSITA would have understood that the other modifications to the bottom branch of the GJBF-CNA are to provide for better tracking performance, which has no impact on the response of the BM or overall characteristics and shape of the response pattern for the bottom branch.

65. For these reasons, it is my opinion that a POSITA would have understood that the overall characteristics and shape of the response patterns would remain unchanged, and therefore would have been motivated to apply Brandstein’s teachings regarding the directivity patterns for the basic GJBF, FBF, and BM to Brandstein’s modified GJBF.

2. Reasonable Expectation of Success

66. In my opinion, a POSITA would have reasonably expected to successfully apply Brandstein’s teachings regarding the frequency insensitivity of the proposed GJBF to the simple GJBF. In my opinion, a POSITA would have recognized that the proposed GJBF operates on the same principles as the simple GJBF (*e.g.*, subtracting the noise-only output of a second virtual microphone from the speech-

and-noise output of a first virtual microphone, to produce an isolated speech signal as the output).

67. In my opinion, the POSITA would have appreciated that, for purposes of isolating the speech signal at the output, the top and bottom branches of both the simple and proposed GJBFs would have maintained the same relevant characteristics, such as those depicted in the Figs. 5.2 and 5.3 directivity plots, including the directivity plots of the final output, FBF, BM, top and bottom branches. Specifically, as shown in Figs. 5.1 and 5.8, the POSITA would have understood that the FBF is the same between the simple and proposed GJBFs (Ex. 1005, 89, 98), such that the response of the top branch (containing both speech and noise) is the same in both.

68. Moreover, despite the structure of bottom branches being different between the simple and proposed GJBFs, the POSITA would have understood that the responses of the bottom branches would have maintained a single null in the speech response toward the direction of speech, as Brandstein depicts for the BM in Fig. 5.3.

69. Brandstein explicitly states that the BM in the GJBF-CNA “provides a wider null for the target with sharper edges.” Ex. 1005, 97. Besides widening the null, the POSITA would have understood that the GJBF-CNA bottom branch does not otherwise modify the overall shape and characteristic of the response pattern

from the BM response shown in Fig. 5.3. This is because the added adaptive filters in the GJBF-CNA bottom branch provide improved tracking to movements of the speech source, but do not change the overall goal of the GJBF to cancel the noise. To cancel the noise, the noise-only signal of the bottom branch must null out the speech, so that the overall shape and characteristics of the BM apply to the bottom branch of the proposed GJBF.

70. Thus, because the entire purpose of the modifications to the bottom branch of the GJBF is to make it more robust to steering vector errors toward the direction of the target without changing the response of the BM elsewhere, and so a POSITA would understand that the BM's response would maintain relevant characteristics, such as having a single null in the speech response toward the direction of speech.

71. For these reasons, it is my opinion that a POSITA would have had a reasonable expectation of success in applying the teachings of Brandstein's GJBF-CNA to Brandstein's simple GJBF.

3. Claim 1

a) [1pre]: "A microphone array comprising:"

72. To the extent the preamble is limiting, it is my opinion that Brandstein discloses this limitation.

73. Brandstein explains that “[a]pplications of beamforming include *microphone arrays* for speech enhancement.” Ex. 1005, 87. In Fig. 5.1, which depicts a basic GSC, the *microphone array* generating microphone signals $x_0, x_1 \dots x_{N-1}$.

b) [1a]: “a first virtual microphone comprising a first combination of a first microphone signal and a second microphone signal,”

74. It is my opinion that Brandstein discloses this limitation.

75. Fig. 5.1 shows *first and second physical microphones* generating *first and second microphone signals* (e.g., x_0 and x_1), respectively. These signals form a *first virtual microphone* along the top branch of the GSC. The *first virtual microphone* is formed from a first combination of the *first and second microphone signals*, using a summer module to combine them (*a first combination*), and delaying the sum:

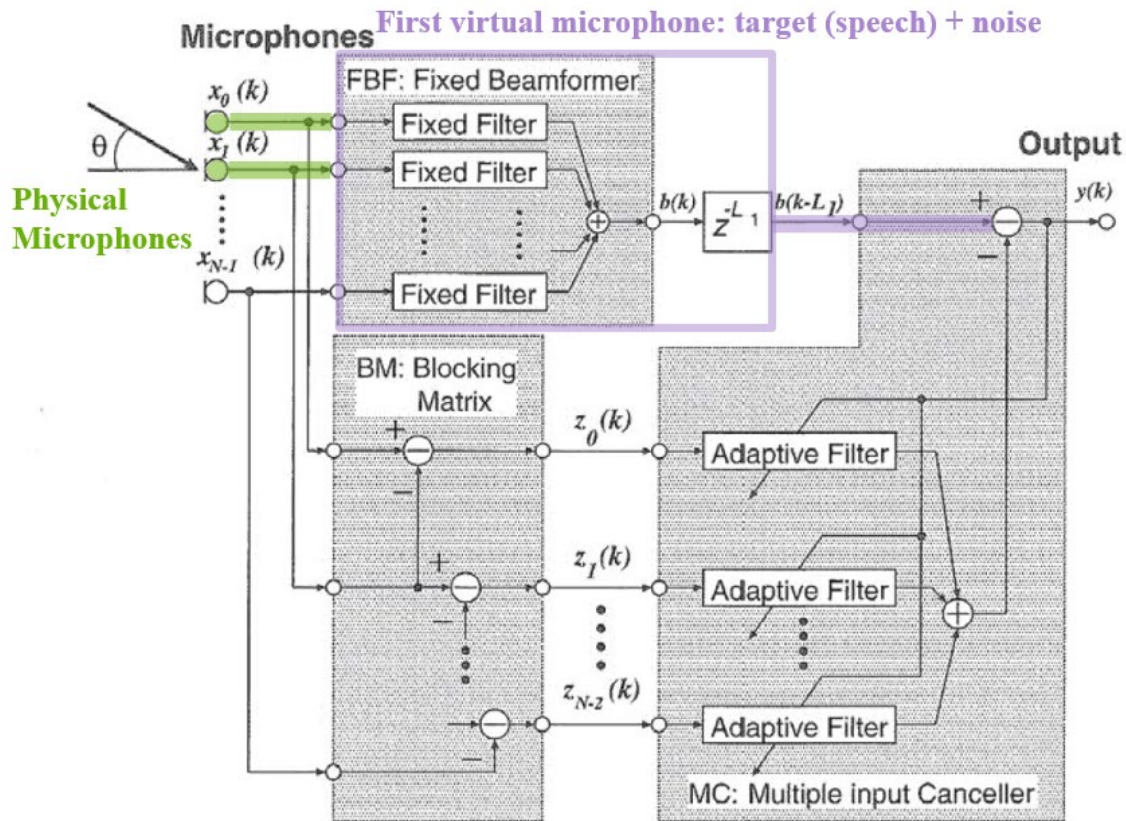


Fig. 5.1. Griffiths-Jim beamformer. It comprises a fixed beamformer (FBF), a multiple-input canceller (MC), and a blocking matrix (BM).

Ex. 1005, Fig. 5.1 (annotated)

Ex. 1005, 88-89, Fig. 5.1.

76. I map the *first virtual microphone* to the GJBF's top branch, which includes an FBF and a delay module. The POSITA would have understood that the response pattern is unaffected by the delay module, so that the response pattern for the FBF is the same as that for the entire top branch.

c) **[1b]: “the first virtual microphone having a first linear response to speech and a first linear response to noise,”**

77. In my opinion, Brandstein discloses this limitation.

78. Brandstein discloses that the *first and second virtual microphones* (top and bottom branches of the GSC, respectively) output separate signals. As discussed in [1a], Brandstein discloses the *first virtual microphone* along the top branch of the GSC that includes an FBF. The top portion of Brandstein’s Fig. 5.3 (reproduced below) shows an example FBF directivity pattern. The horizontal axis is direction of arrival measured in degrees relative to the microphone array. The target signal is *speech*, which is shown at 0 degrees, while the interference signal, which is *noise*, is positioned at approximately 45 degrees.

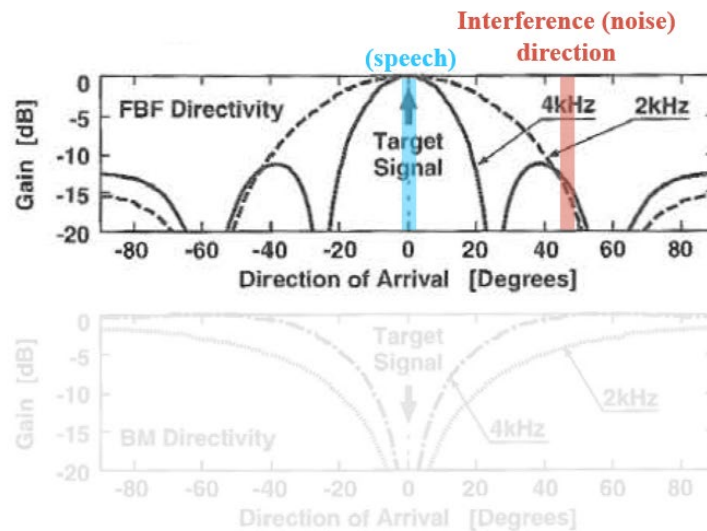


Fig. 5.3. Directivity pattern of a fixed beamformer (FBF) and a blocking matrix (BM).

Ex. 1005, Fig. 5.3 (annotated)

79. The '691 patent defines “noise” as “unwanted environmental acoustic noise.” Ex. 1001, 5:40-41. In my opinion, a POSITA would understand that Brandstein refers to environmental acoustic noise as “interference” because Brandstein states that interference can be, for example, air-conditioner noise or white noise. Ex. 1005, 101, 103 Fig. 5.13.

80. Brandstein discloses that the target signal is *speech* because it states that “[t]he target source was an English male speech signal.” *E.g.*, Ex. 1005, 101; *also id.*, 103, Figs. 5.13, 5.14. In noting that the GSC was the most widely known

adaptive beamformer, Brandstein states: “Applications of beamforming include microphone arrays for speech enhancement. The goal of speech enhancement is to remove undesirable signals such as noise and reverberation.” Ex. 1005, 87.

81. As shown in Fig. 5.3, the gain (in dB) for the FBF of the top branch (*first virtual microphone*) has *a first linear response to speech* in the direction of a “target signal” (*speech*) and a *first linear response to interference, or noise*, in the direction of the interference. The FBF’s gain is a *response* of the *first virtual microphone* because the delay module has no effect on the directivity pattern of the beamformer.

82. As I explained in Section VI.C., Brandstein subtracts the output of the GSC’s bottom branch from the output of the top branch. Fig. 5.2 shows the directivity pattern for the resulting output of the GJBF, with the same horizontal axis as in Fig. 5.3:

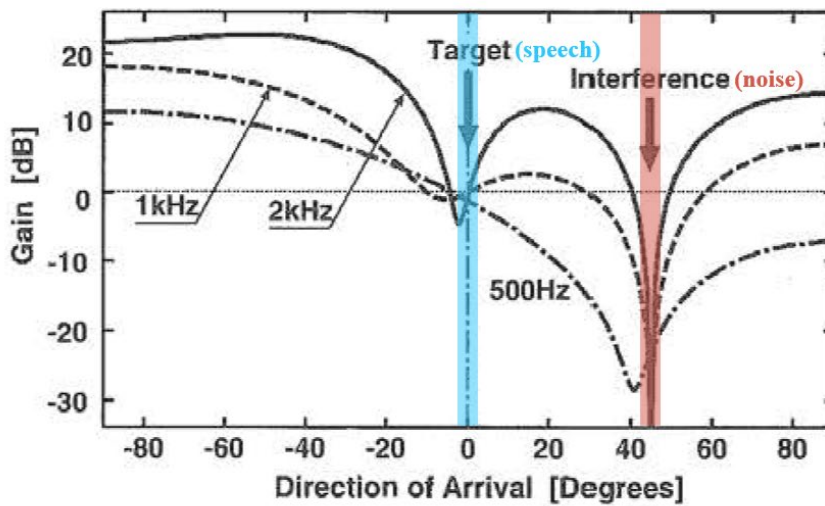


Fig. 5.2. Example directivity pattern of the Griffiths-Jim beamformer.

Ex. 1005, Fig. 5.2 (annotated)

Fig. 5.2 shows that “in the direction of the interference, a deep null is formed.” Ex. 1005, 90. This “deep null” in the direction of the interference, or *noise*, is formed by subtracting the linear response of the bottom branch (second virtual microphone) in the direction of *noise* from the linear response of the top branch (*first microphone*) in the direction of *noise*. Therefore, the *first virtual microphone* of the GSC has a *first response to speech* and a *first response to noise*.

83. A POSITA would have recognized that the *responses* to speech and noise of Brandstein’s first virtual microphone’s are *linear* because the responses are represented by a linear system. For example, Brandstein’s Fig. 5.1 depicts the top

branch as applying “fixed filter[s]” that are summed together followed by a delay module, which are all *linear* components. Ex. 1005, 95, Figs. 5.1, 5.3.

- d) **[1c]: “the first linear response to speech being substantially similar across a plurality of frequencies for a speech source located at an angle relative to an axis of the microphone array,”**

84. In my opinion, Brandstein renders this limitation obvious.

85. As I explained in VIII.A.1., a POSITA would have been motivated to apply the teachings of Brandstein’s GJBF-CNA to modify its basic GJBF, while understanding that the overall shape and characteristics of the response patterns of the basic GJBF remain unchanged for the modified GJBF. For the simple GJBF, Brandstein depicts the overall shapes of the directivity patterns for a limited set of frequencies for the FBF (in the top branch) and BM (in the bottom branch with sharp null) (*see* Fig. 5.3) as well as for final output (*e.g.*, the difference between top and bottom branches (*see* Fig. 5.2)). A POSITA would have recognized that Brandstein’s teachings about these directivity patterns are unchanged for the modified GJBF.

86. Fig. 5.10 of Brandstein shows a plot of the sensitivities to the frequency component of the target signal (*first linear response to speech across a plurality of frequencies*) at 500, 1000, 2000, 3000, and 3500 Hz for this CNA-modified GSC:

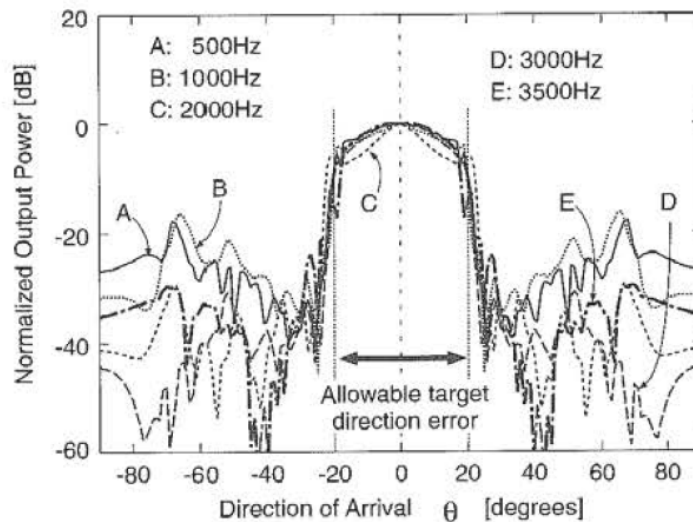


Fig. 5.10. Sensitivities after convergence as a function of DOA at different frequencies.

87. Brandstein explains that the “[f]requency dependency of the GJBF-CNA is small” and that the sensitivity in the region around the target direction (i.e. direction of speech) “is constant.” Ex. 1005, 99-100. In my opinion, this indicates that the response to speech of the GJBF-CNA is similar across a plurality of frequencies.

88. The POSITA would have understood that Fig. 5.10 depicts the output $y(k)$ of Fig. 5.8, as Brandstein describes the plot in relation to “the GJBF-CNA.” Ex. 1005, 100. Moreover, because the output of Brandstein’s bottom branch in Fig. 5.8 has a null in the direction of speech because it has “a wider null for the target [speech] with sharper edges,” the POSITA would have further understood that the

Fig. 5.10 plot in the target *speech* direction would have also been the top branch's response in the same *speech* direction, at different *frequencies*. For these reasons, the top branch's response to *speech* of the GJBF-CNA (*first response*) is constant (which meets *substantially similar*) across a plurality frequencies.

89. Thus, Brandstein teaches that the directivity pattern of the top branch for speech (*first linear response to speech*) is *substantially similar across a plurality of frequencies for a speech source* because the plots for different frequencies in Fig. 5.10 are nearly identical around the direction of speech. In Brandstein, the range of angles for which the plots are nearly identical is $\theta = -20^\circ$ to 20° , where θ is *an angle relative to an axis of the microphone array* as shown in Fig. 5.1.

- e) [1d]: “wherein the first microphone signal is generated by a first physical microphone and the second microphone signal is generated by a second physical microphone; and”

90. In my opinion, Brandstein discloses this limitation.

91. Fig. 5.1 of Brandstein shows that *the first microphone signal, x_0 , is generated by a first physical microphone*, and that *the second microphone signal, x_1 , is generated by a second physical microphone*.

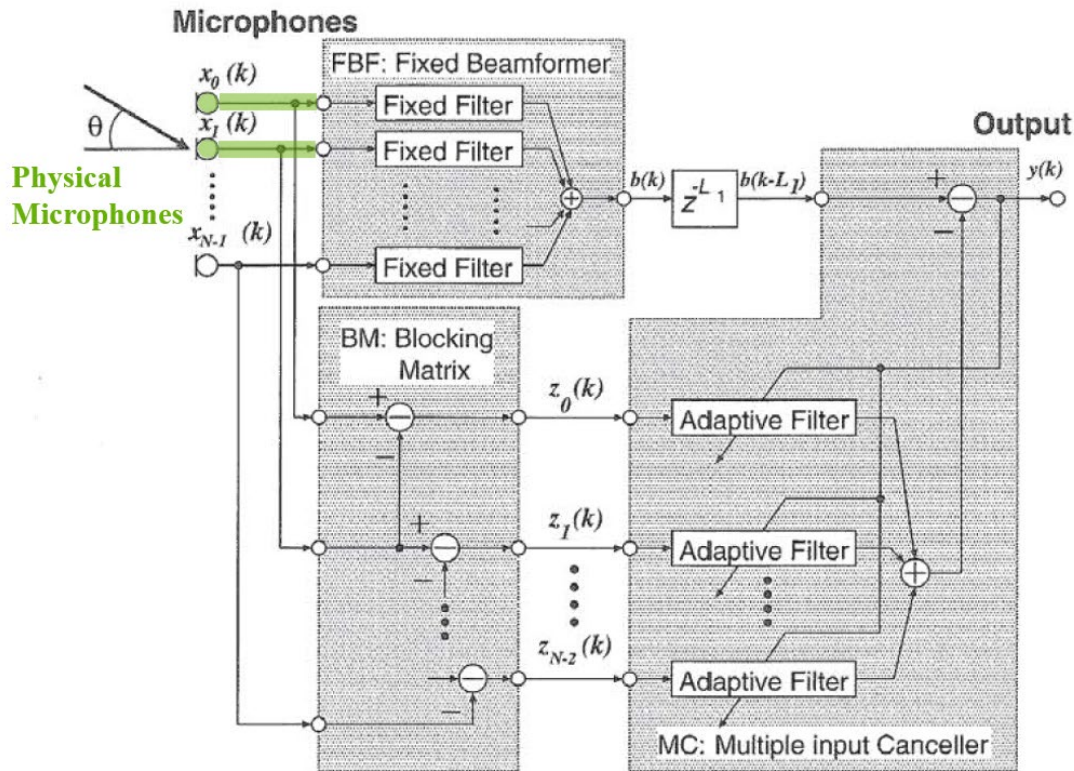


Fig. 5.1. Griffiths-Jim beamformer. It comprises a fixed beamformer (FBF), a multiple-input canceller (MC), and a blocking matrix (BM).

Ex. 1005, Fig. 5.1 (annotated)

- f) [1e]: “a second virtual microphone comprising a second combination of the first microphone signal and the second microphone signal,”

92. In my opinion, Brandstein discloses this limitation.

93. Fig. 5.1 shows a *second virtual microphone* along the bottom branch of the GSC that is formed from a *second combination* of the *first and second microphone signals* (e.g., x_0 and x_1). The *second combination* is different from the combination in the top branch, because the bottom branch is formed by a BM, in which

one microphone signal is subtracted from another, followed by adaptive filtering in the MC, while the top branch is formed by summing the filtered outputs in the FBF:

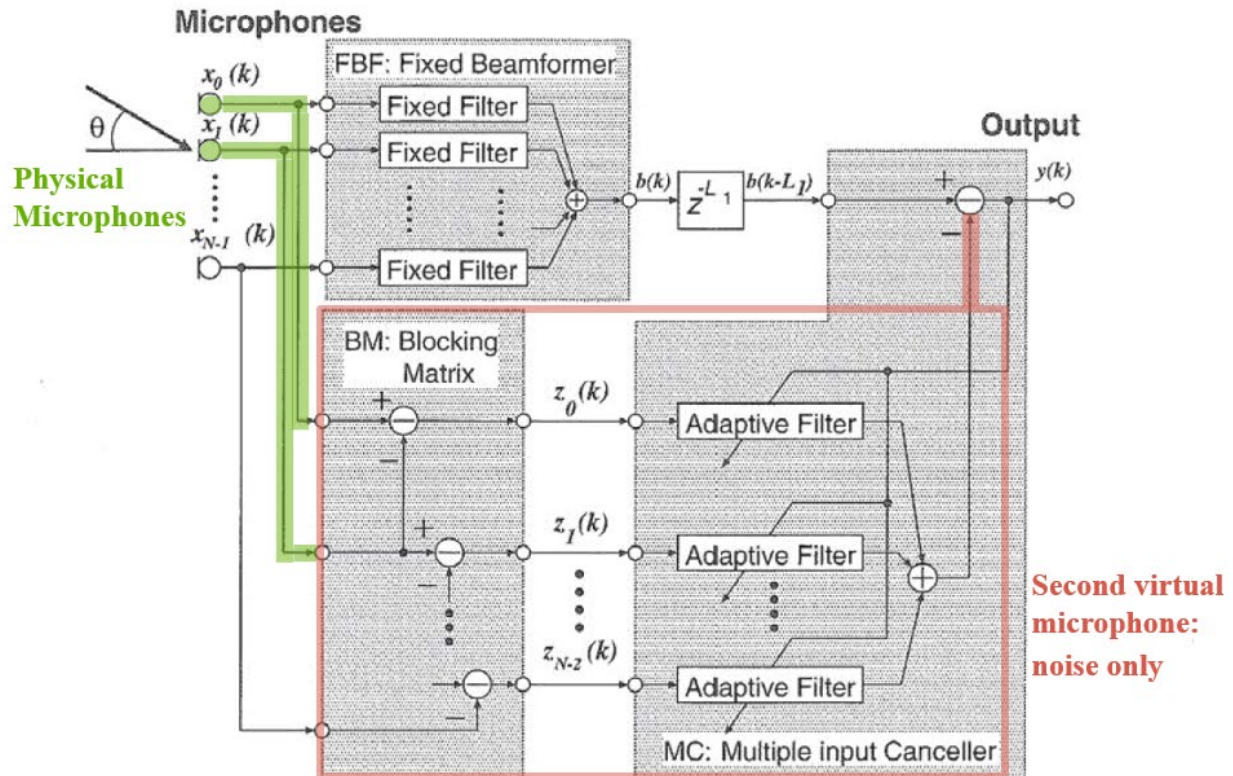


Fig. 5.1. Griffiths-Jim beamformer. It comprises a fixed beamformer (FBF), a multiple-input canceler (MC), and a blocking matrix (BM).

Ex. 1005, Fig. 5.1 (annotated)

Ex. 1005, 89, Fig. 5.1.

94. I map the *second virtual microphone* to the GJBF's bottom branch, which includes a BM and adaptive filters and a summer module in the MC. In my opinion, the POSITA would have understood that the response pattern for the BM is

the same as that for the entire bottom branch, as its overall shape and characteristics are not significantly affected by the MC components.

- g) [1f]: “the second virtual microphone having a second linear response to speech and a second linear response to noise,”

95. In my opinion, Brandstein discloses this limitation.

96. As depicted in Brandstein’s Fig. 5.3, the BM of the bottom branch (*second virtual microphone*) has *a second linear response to speech*, which is in the direction of a “target signal” (*speech*):

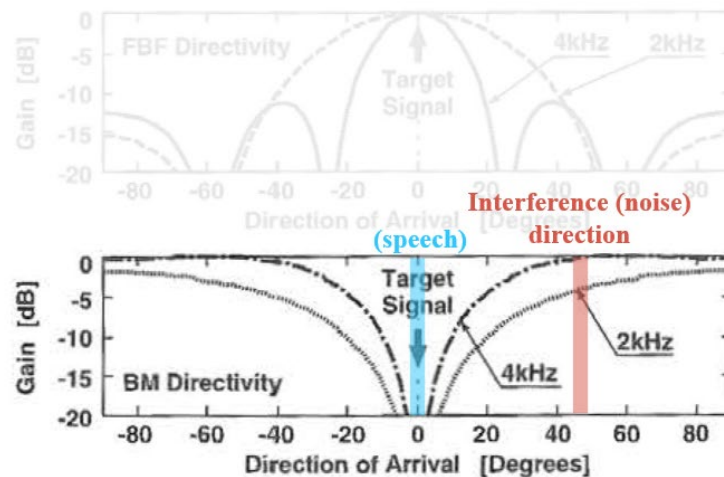


Fig. 5.3. Directivity pattern of a fixed beamformer (FBF) and a blocking matrix (BM).

Ex. 1005, Fig. 5.3 (annotated)

97. For the same reasons as discussed above with respect to [1b], the bottom branch of the GSC (*second virtual microphone*) also has *a second response to noise*. The second virtual microphone's response to the speech and response to the noise are *linear responses* because the bottom branch is represented by difference modules, adaptive filters, and a summer, which are all *linear* components.

98. Although the bottom of Fig. 5.3 depicts the BM response, which only forms part of the bottom branch. Ex. 1005, 88. As Brandstein explains, "the BM forms a null in the look direction so that the target signal is suppressed and all other signals are passed through." Ex. 1005, 88. The MC performs additional processing with adaptive filters that "generate replicas of components correlated with the interferences." The POSITA would have known that the general characteristics of the bottom branch (*e.g.*, the output of the MC) would have also had a *second linear response speech* and a *second linear response to noise* similar to those of the BM, as the main purpose of the bottom branch is to isolate noise for subtraction at the output and because the MC's filters would not significantly change the directivity patten of the BM response.

- h) **[1g]: "the second linear response to noise being substantially similar to the first linear response to noise, and the second linear response to speech being**

substantially dissimilar to the first linear response to speech,”

99. In my opinion, Brandstein discloses or renders obvious this limitation.

100. It is my opinion that a POSITA would have known that Brandstein’s virtual microphones have very different (which meets the *substantially dissimilar* limitation) responses to the target signal (*speech*) because the first virtual microphone is designed to form a beam that captures the target signal, while the second virtual microphone is designed to block the target signal. Ex. 1005, 88.

101. A POSITA also would have known that the two virtual microphones also have essentially the same (which meets the *substantially similar* limitation) responses to *noise* so that in the final subtraction output the noise cancels out. Ex. 1005, 89, Fig. 5.2.

102. The very different responses to *speech* and nearly identical responses to *noise* are evident from Fig. 5.2, which shows the directivity pattern for the final output of the simple GJBF, and Fig. 5.3, which shows the directivity patterns of the FBF and BM:

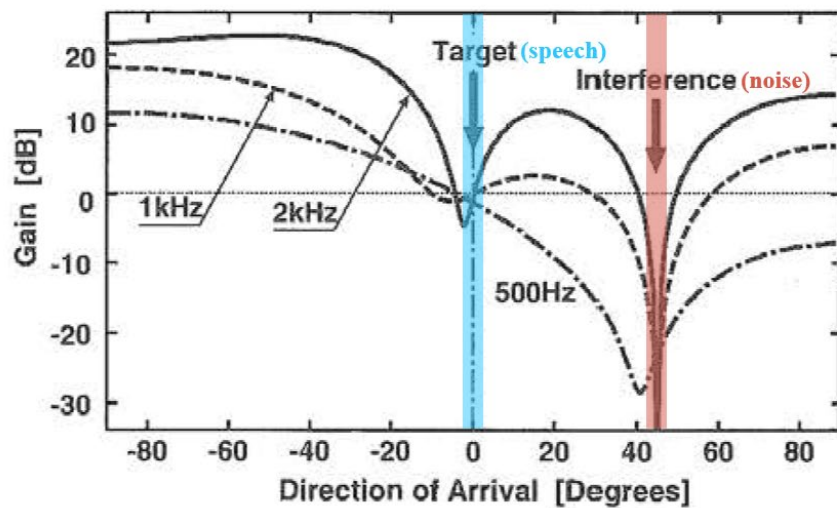


Fig. 5.2. Example directivity pattern of the Griffiths-Jim beamformer.

Ex. 1005, Fig. 5.2 (annotated)

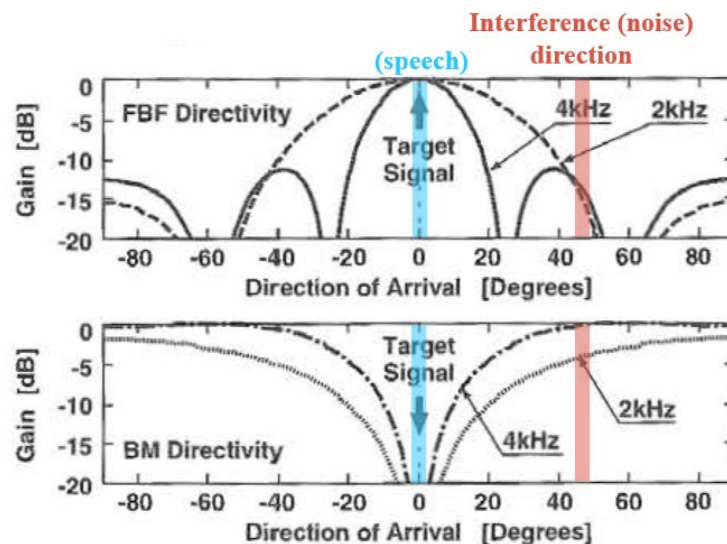


Fig. 5.3. Directivity pattern of a fixed beamformer (FBF) and a blocking matrix (BM).

Ex. 1005, Fig. 5.3 (annotated)

Ex. 1005, 89.

103. Brandstein states that in the direction of the target signal (highlighted in blue), the GJBF achieves “almost constant gains close to 0 dB.” Ex. 1005, 90. A POSITA would have known that a gain of zero means that subtracting the second virtual microphone’s noise-only output from the first virtual microphone’s speech-and-noise output would produce the target signal in the direction of the target, thereby showing that the two virtual microphones have very different responses to the target signal. Therefore, in my opinion, this shows *the second linear response to speech being substantially dissimilar to the first linear response to speech* limitation. This is also evident in Fig. 5.3, which shows that the FBF’s response in the direction of speech has a maximum, which means it is very different from the BM’s response in the direction of speech having a null.

104. In contrast, Brandstein states that “in the direction of the interference, a deep null is formed” as highlighted in red above. Ex. 1005, 90. A POSITA would have known that the deep null reflects that subtracting the two virtual microphones’ outputs cancels the interference, and confirms that the virtual microphones’ responses to the interference are essentially the same. Therefore, in my opinion, this shows that *the second linear response to noise being substantially similar to the first linear response to noise* limitation.

105. The GJBF-CNA also meets this limitation, as Fig. 5.10 depicts a gain of 0 dB at zero degrees for the target *speech* signal and a very negative gain at 45 degrees for the interference *noise* signal. Ex. 1005, 100.

- i) **[1h]: “wherein at least one of the first linear response to noise and the second linear response to noise is being non-zero in a direction towards a source of noise.”**

106. In my opinion, Brandstein discloses this limitation.

107. As shown in Brandstein’s Fig. 5.3, the only null in the directivity pattern of the BM of the bottom branch of the GJBF (*second virtual microphone*) is in the direction of speech. In other words, the response in all other directions is non-zero, including *in a direction toward the source of noise*. Therefore, at least the *second linear response to noise is nonzero in a direction toward the source of noise*.

4. Claim 2: “The microphone array of claim 1, wherein the first and second physical microphones are omnidirectional.”

108. In my opinion, Brandstein discloses this limitation.

109. Brandstein explains that omni-directional microphones can be used in the GSC described. Ex. 1005, 101. A POSITA have understood that omni-directional microphones can be used in either the basic GSC described in Fig. 5.1 or the GJBF-CNA described in Fig. 5.8.

5. Claim 3: “The microphone array of claim 1, wherein the first linear response to speech is devoid of a null, wherein the speech is human speech.”

110. In my opinion, Brandstein discloses or renders obvious this limitation.

111. As shown in the top directivity plot of Brandstein Fig. 5.3 for the simple GJBF (*the first linear response to speech*), there is no zero or minima in the spatial response in the direction of the target. Therefore, the simple GJBF has a *first linear response to speech* that is *devoid of a null*.

112. The sensitivity plot of Fig. 5.10 also shows that there is no null in the target (*speech*) direction because there is no zero or minima in the directivity pattern (*response*) of the FBF (part of the *first virtual microphone*) in the target direction. Therefore, Brandstein discloses a *first linear response to speech* that is *devoid of a null*.

113. Alternatively, a POSITA would have understood that the directivity pattern (*first linear response to speech*) of the FBF is *devoid of a null* in all directions for certain frequencies. A POSITA would have seen from the top directivity plot in Fig. 5.3 that the main lobe with the largest maximum in the directivity is centered around the direction of speech (DOA of 0°) and would have understood that the width of this maximum depends on frequency. Specifically, a POSITA would have seen that the width of the maximum approximately doubles from approximately 50°

to 100° when the frequency is approximately halved from 4 kHz to 2 kHz. Therefore, a POSITA would have recognized that for a frequency of below about 1000 Hz, the main lobe's width would span an entire range of directions plotted in Fig. 5.3, such that the *response is devoid of a null* in any direction covered by the plot. Moreover, even for the directions not depicted in the top directivity plot in Fig. 5.3, it is my opinion that a POSITA would have understood, based on these same observations, that the directivity pattern (*first linear response to speech*) of the FBF is *devoid of a null* in all directions.

114. Brandstein also teaches that the *speech is human speech*, by disclosing that the target speech signal is “an English male speech signal.” Ex. 1005, 101.

6. Claim 4: “The microphone array of claim 3, wherein the second linear response to speech of the second virtual microphone includes a single null oriented in a direction toward a source of the speech.”

115. In my opinion, Brandstein discloses this limitation.

116. Fig. 5.3 shows that the BM of the bottom branch (*second virtual microphone*) has a *second linear response to speech*, which is in the direction of a “target signal” (*speech*):

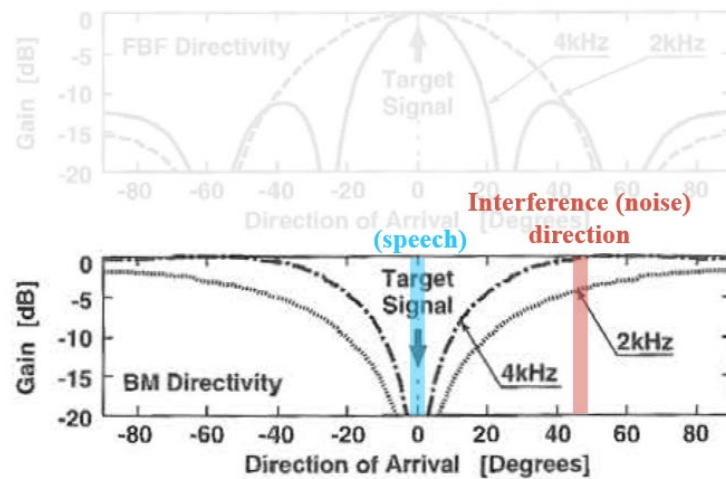


Fig. 5.3. Directivity pattern of a fixed beamformer (FBF) and a blocking matrix (BM).

Ex. 1005, Fig. 5.3 (annotated)

117. The general characteristics and overall shape of the directivity pattern of the BM are the same as that of the *second virtual microphone* because the purpose of both is to isolate the *noise* for subtraction at the output. As shown in Fig. 5.3, in the direction of the target speech signal, the BM has *a single null oriented in the direction toward the direction of the target signal (a source of speech)*.

118. Additionally, if the *response to speech* is not limited to the direction of the target *speech* and includes a wider range of directions, a POSITA would have understood the response of the bottom branch (*second linear response to speech*)

has a *single null oriented in a direction toward a source of the speech* across different directions. First, the bottom plot of Figure 5.3 depicts on a single null for the directions ranging from -90 to 90 degrees.

119. Additionally, although Brandstein's Figure 5.3 does not expressly show the response for a full 360 degrees, a POSITA would have known that the response of the BM has a *single null oriented in a direction toward a source of the speech* for certain speech directions. For example, this is the case if the speech direction is parallel to the axis of the microphone array. The POSITA's understanding is reflected in, for example, Figs. 3 and 4a of U.S. Patent No. 7,386,135 ("Fan," Ex. 1012.

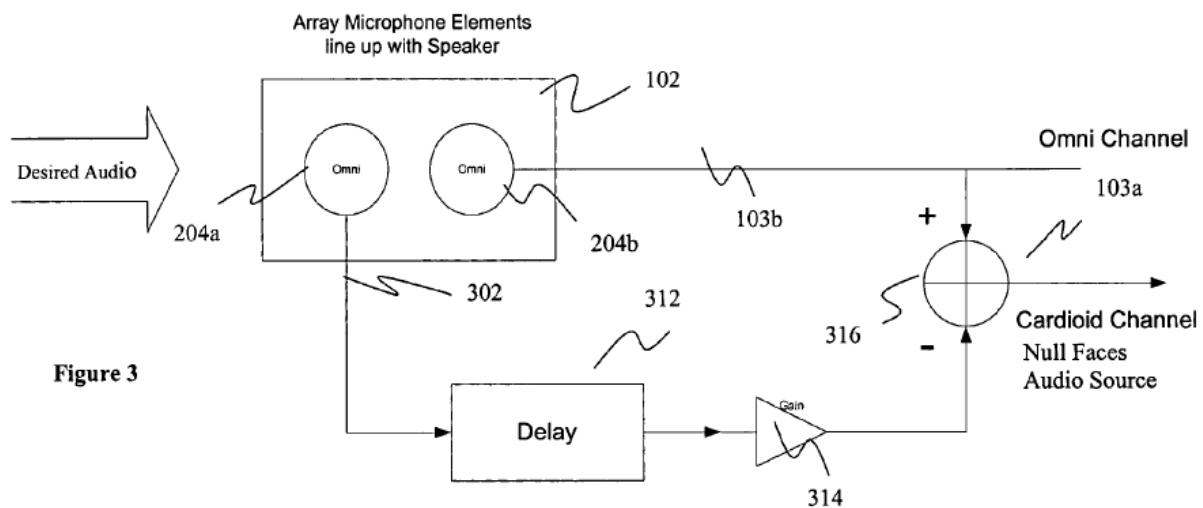
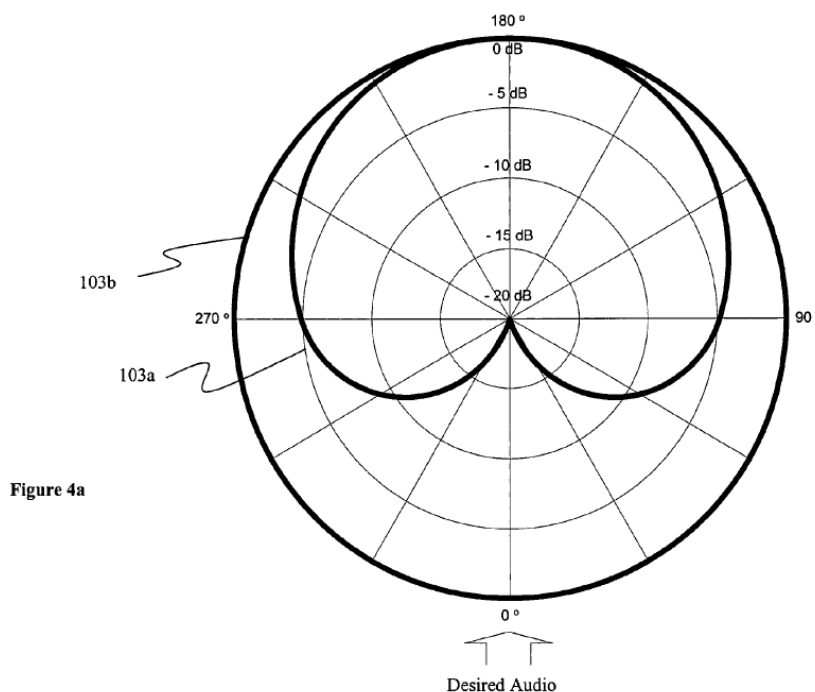


Fig. 3 depicts a delay-and-subtract structure like that of the GJBF's BM. The signal from the omnidirectional physical microphone 204a is delayed and then subtracted from the signal from the other omnidirectional physical microphone 204b.

120. Fig. 4a “illustrates the acoustic beams generated by the various acoustic devices of FIG. 2a-2c, and FIG.2g (when supplemented with the circuitry of FIG.3) in the form of a polar sensitivity plots.” Fig. 2g depicts the 2-omindirectional microphone array 102. Fig. 4a, reproduced below, shows that the response of Fig. 3's delay-and-subtract structure has *a single null oriented in a direction toward the “desired audio” (source of the speech)*. Ex. 1012, 2:16-23 (emphasis added), Fig. 4a.



7. **Claim 5: “The microphone array of claim 4, wherein the single null is a region of the second linear response to speech having a measured response level that is lower than the measured response level of any other region of the second linear response to speech.”**

121. In my opinion, Brandstein discloses this limitation.

122. As shown in the bottom plot of Fig. 5.3, the *single null* has a gain (*measured response level*) that more negative (*lower than the measured response level*) compared to that of *any other region* of the directivity pattern, which includes the *second linear response to speech*. This includes any region away from zero degrees.

8. **Claim 6: “The microphone array of claim 4, wherein the second linear response to speech includes a primary lobe oriented in a direction away from the source of the speech.”**

123. In my opinion, Brandstein discloses or renders obvious this limitation.

124. As explained above for claims 4-5, the second linear response to speech includes a null in the direction of speech. According to the '691 patent, the primary lobe “is a region of the second linear response having a measured response level that is greater than the measured response level of any other region of the second linear response.” (Ex. 1001, claim 7, 19:4-7, 20:21-24, 21:16-19, 22:4-7, 22:45-48, 24:37-40, 25:6-10, 25:53-56, 27:21-24, 28:11-14, 29:20-23, 30:8-11, 31:45-48, 32:18-21, 32:55-58.)

125. As the bottom directivity pattern of Fig. 5.3 shows, the gain (*response*) increases as the direction of arrival extends away from the *speech* direction at 0° (within the blue region in the annotated Fig. 5.3, above for [1f] and claim 4). Therefore, the *second linear response to speech includes a primary lobe oriented in a direction away from the source of the speech.*

126. Additionally, if the *response to speech* is not limited to the direction of the target *speech* and includes any direction, it is my opinion that Brandstein renders obvious this limitation. Because the bottom plot of Fig. 5.3 depicts a null in the speech direction, a POSITA would have understood that all lobes in the directivity plot are oriented in a direction away from the source of the speech, including a primary lob. Thus, a POSITA would have understood the response of the BM has *includes a primary lobe oriented in a direction away from the source of the speech,* for certain speech directions. For the same reasons as explained above for claim 4, if the speech direction is parallel to the axis of the microphone array, then the *second linear response* to speech would have the shape depicted in Fan Fig. 4a, having a single null towards speech and nonzero everywhere else. Ex. 1012. Fig. 4a shows that the response of Fig. 3's delay-and-subtract structure has *a primary lobe oriented in a direction away from the direction of the "desired audio" (source of the speech),*

because the lobe is oriented exactly opposite to the speech direction. Ex. 1012, 2:16-23, Fig. 4a.

9. **Claim 7: “The microphone array of claim 6, wherein the primary lobe is a region of the second linear response to speech having a measured response level that is greater than the measured response level of any other region of the second linear response to speech.”**

127. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above with respect to claim 6.

10. **Claim 8: “The microphone array of claim 4, wherein the first physical microphone and the second physical microphone are positioned along an axis and separated by a first distance.”**

128. In my opinion, Brandstein discloses this limitation.

129. In Fig. 5.1 of Brandstein, which shows the generalized structure for a GSC, and Fig. 5.13, which shows an exemplary implementation of the GSC, the first and second physical microphones are *positioned along an axis* that is denoted below by the dotted line:

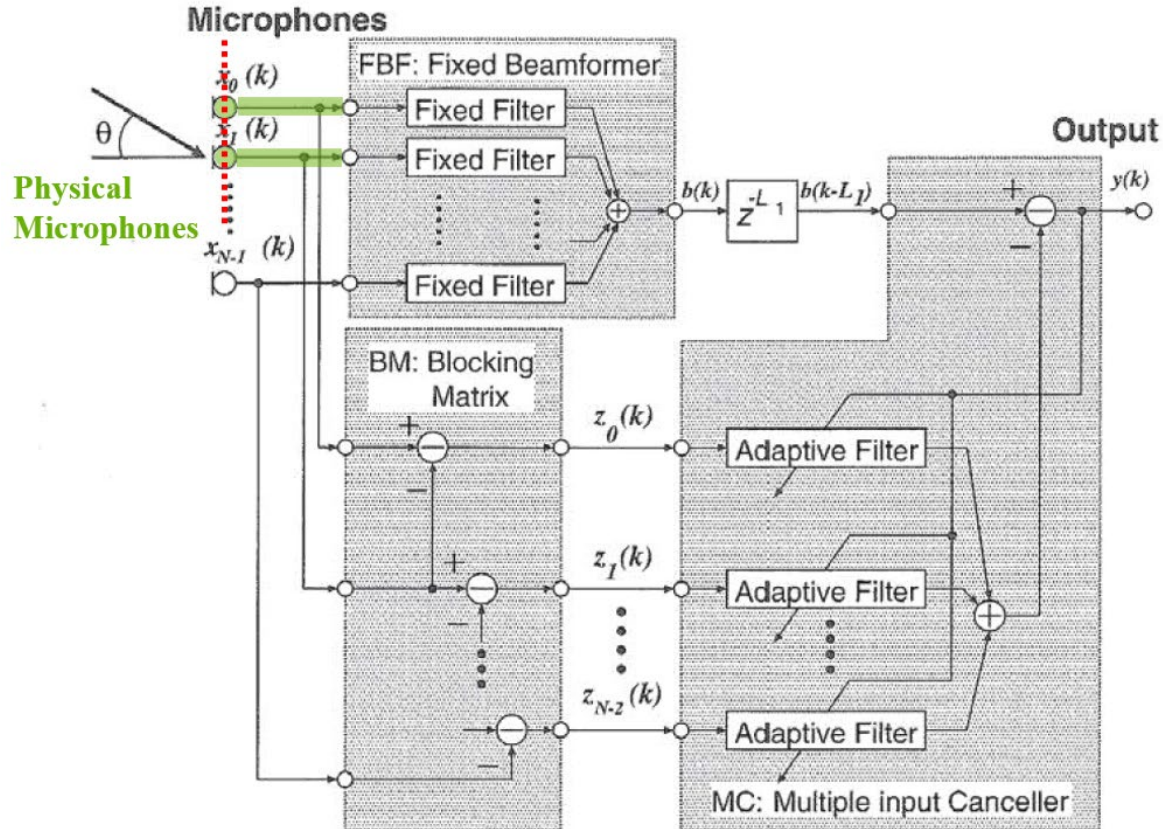


Fig. 5.1. Griffiths-Jim beamformer. It comprises a fixed beamformer (FBF), a multiple-input canceler (MC), and a blocking matrix (BM).

Ex. 1005, Fig. 5.1 (annotated)

130. This is similarly shown in Fig. 5.13, which shows an exemplary implementation of the GSC:

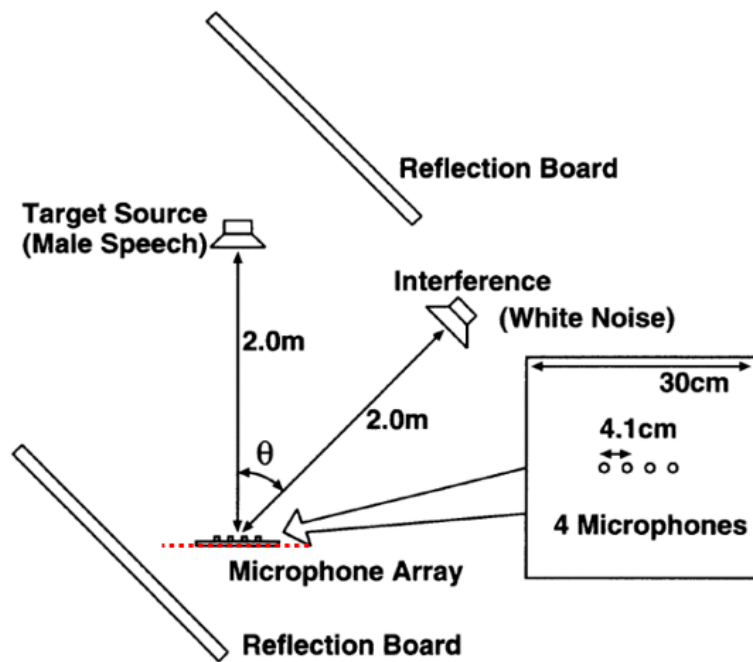


Fig. 5.13. Experimental set-up.

Ex. 1005, Fig. 5.13 (annotated)

As shown, each physical microphone in the GSC is separated from each other by a distance, including the first and second physical microphones. Brandstein teaches an exemplary implementation where omni-directional physical microphones are equally spaced at 4.1 cm intervals, and therefore shows that the *physical microphone are positioned along an axis and separated by a first distance* (e.g., 4.1 cm). Ex. 1005, 101, Fig. 5.13.

11. Claim 9: “The microphone array of claim 8, wherein a midpoint of the axis is a second distance from the speech source that generates the speech, wherein the speech source is located in a direction defined by an angle relative to the midpoint.”

131. In my opinion, Brandstein discloses or renders obvious this limitation.

132. As explained for claim 8, Brandstein discloses that the microphone units are placed on an axis at a certain distance apart from each other. A midpoint of the axis is located halfway between the microphones along that axis. This midpoint is necessarily located at a distance (*second distance*) *from the speech source that generates the speech*. A POSITA would have understood that the speech source is located in a direction defined by an angle relative to the midpoint of the axis. For example, Brandstein teaches an exemplary implementation of the GSC where the male speech target source (*speech source*) is *located in a direction relative to the midpoint*, perpendicular or 90° from the midpoint, and at a distance (second distance from the speech source that generates the speech) 2.0 m away. Ex. 1005, 103, Fig. 5.13. A POSITA would further have known that although Fig. 5.13 shows the target source as being perpendicular, this is not required and the target source can be in any direction. If the target direction is not perpendicular, then a delay element is added. Ex. 1005, 88 (“Assuming a look direction perpendicular to the array surface, no delay element is necessary.”).

12. Claim 10: “The microphone array of claim 9, wherein the first virtual microphone comprises the second microphone signal subtracted from the first microphone signal.”

133. In my opinion, Brandstein discloses or renders obvious this limitation.

134. The *first microphone signal* and the *second microphone signal* are input into filters in the top branch (*first virtual microphone*) of the GSC and then summed.

135. A POSITA would have understood that the filters apply weight coefficients (representing the filter’s impulse response including positive and/or negative values) to the incoming microphone signals through a mathematical convolution, and that these weight coefficients typically include both positive and negative weights.

136. In a convolution, the weight coefficients are multiplied by samples of the microphone signals as they are received over time. Thuse, applying a negative weight to the *second microphone signal* means that the filtered output includes a negative version of the *second microphone signal* such that the FBF’s summer would *subtract* it from the *first microphone signal*. Furthermore, in such a convolution, the filters are implemented by multiplying the weights times *delayed* versions of each *microphone signal*.

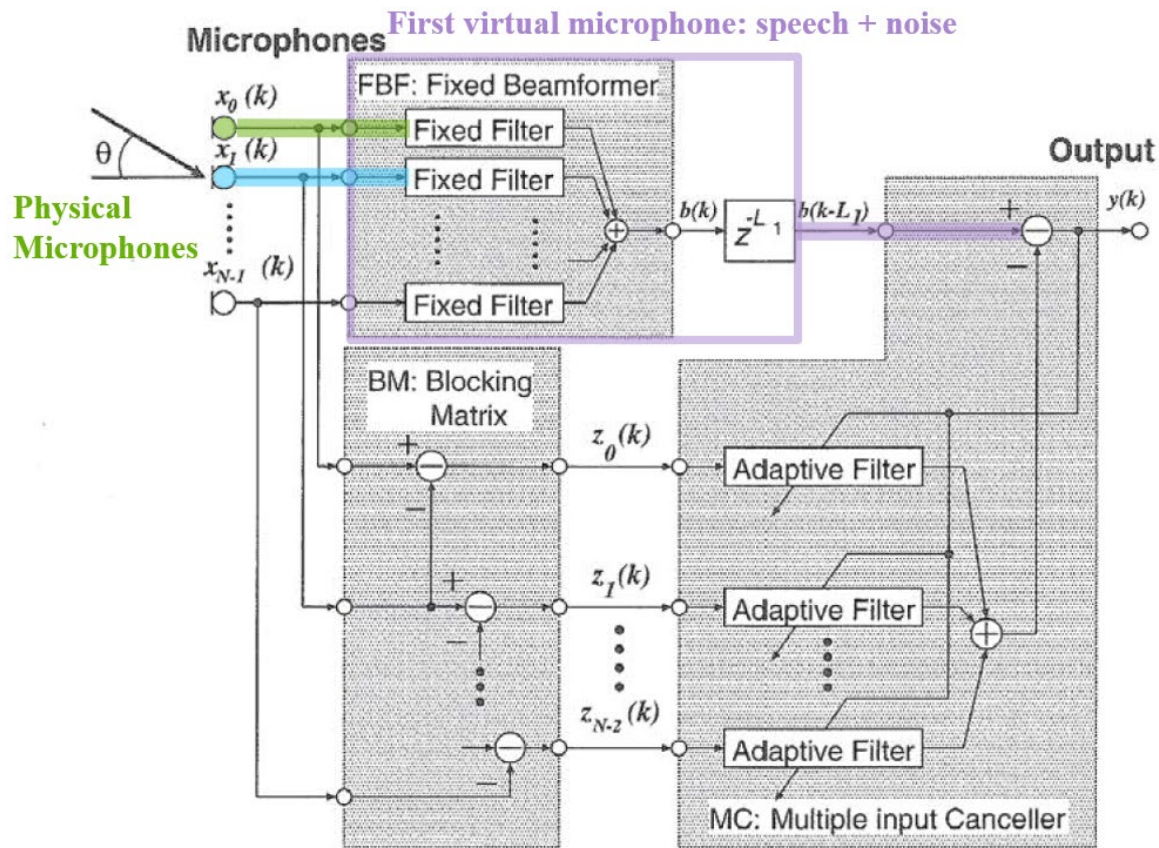


Fig. 5.1. Griffiths-Jim beamformer. It comprises a fixed beamformer (FBF), a multiple-input canceller (MC), and a blocking matrix (BM).

Ex. 1005, 89, Fig. 5.1 (annotated)

137. As an illustration, the weight coefficients for an example simple low-pass filter structure for the first two physical microphones are shown below. The coefficients are multiplied times a delayed copy of the input signal, then summed to get the overall output of the FBF. For these example filter coefficients, the individual filter outputs and total signal labeled “ $b(k)$ ” in the above diagram, can be written as follows:

$$\begin{aligned} b(k) = & +0.354x_0[k-19]-0.455x_0[k-17]+0.0637x_0[k-15]-0.1061x_0[k-13]+0.3183x_0[k-11] \\ & +0.5x_0[k-10]+0.3183x_0[k-9]-0.1061x_0[k-7] +0.0637x_0[k-5]-0.455x_0[k-3]+0.354x_0[k-1] \\ & +0.318x_1[k-20]-0.250x_1[k-19]-0.322x_1[k-17]-0.0531x_1[k-16] \\ & -0.045x_1[k-15]+0.075x_1[k-13]+0.1592x_1[k-12]+0.2251x_1[k-11] \\ & +0.25x_1[k-10] +0.2251x_1[k-9]+0.1592x_1[k-18]+0.075x_1[k-7]-0.045x_1[k-5] \\ & -0.0531x_1[k-4]-0.322x_1[k-3]-0.250x_1[k-1]+0.318x_1[k-] \end{aligned}$$

138. As shown in this example, a POSITA would understand that in this configuration, the output of the FBF, “ $b(k)$ ”, *comprises the second microphone signal* (“ x_1 ”) *subtracted from a delayed version of the first microphone signal* (“ x_0 ”).

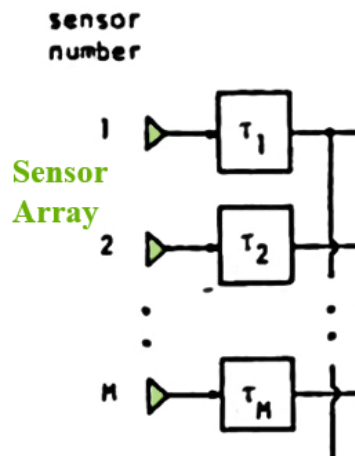
139. Moreover, the POSITA would have understood that the labeling of Brandstein’s microphone signals x_0 and x_1 as the *first and second microphone signals* (respectively) could be flipped, without affecting the analysis because (1) all of the requirements in the claims from which claim 10 depends (claims 1, 3, 4, 8, 9) with respect to both *microphone signals* are identical and apply to both *microphone signals* symmetrically and (2) the mapping for the limitation of claim 10 to applying negative weights to a *microphone signal* would apply to either *microphone signal*. Ex. 1005, 89, Fig. 5.1 (annotated with an alternative mapping).

13. Claim 11: “The microphone array of claim 10, wherein the first microphone signal is delayed.”

140. In my opinion, Brandstein discloses or renders obvious this limitation.

141. Brandstein explains that the “BM [i.e., blocking matrix] in Fig. 5.1 has a simple delay-and-sum structure” which in the case of the BM is also called a “delay-and-subtract” structure since the signals are subtracted (summed with a negative sign) to remove the target signal. Ex. 1005, 88-90. In the case illustrated in Fig. 5.1, the target signal’s assumed direction of arrival (the “look direction”) is perpendicular to the array surface, so “no delay element is necessary.” Ex. 1005, 88-89. Therefore, when θ in Fig. 5.1 is zero, the BM’s delay-and-subtract structure can simplify to “a set of subtracters” Ex. 1005, 88. Brandstein explains that “[t]his structure is actually the one shown in Fig. 5.1.” Ex. 1005, 88. In my opinion (and as Brandstein implies) a POSITA would have understood that, when the target signal is expected to arrive from a non-perpendicular direction (i.e., when θ in Fig. 5.1 is nonzero, as Fig. 5.1 itself indicates is possible), the delay-and-subtract structure is not simplified. Instead, before subtracting, delays are used to time-align the target-signal components of the microphone signals. Indeed, Brandstein says that “[d]elay-and-sum beamformers apply time shifts to the array signals to compensate for the propagation delays in the arrival of the source signal at each microphone. These signals are time-aligned and summed together to form a single output signal.” This time alignment is what allows the subtraction operations to block the target signal.

142. The time alignment implicit in Brandstein's reference to a "delay-and-subtract" structure for the BM is expressly laid out in the original Griffiths-Jim paper:



Ex. 1006, 30, Fig. 4 (excerpted and annotated).

The elements labeled $\tau_1, \dots, \tau_M$ are delays used to time-align the sensor signals. Ex. 1006.

143. Griffiths-Jim therefore expressly shows aligning the sensor signals before feeding them to the BM, just as Brandstein contemplates, and just as a POSITA would have understood. Indeed, Brandstein presents Fig. 5.1 and the related discussion precisely to explain the original Griffiths-Jim paper. Ex. 1005, 88-90, 108. A POSITA reading Brandstein would have thus immediately understood that Brandstein's reference to a "delay-and-subtract" structure for the BM means the signals

should be aligned in time, which can be addressed through conventional delays of the kind Griffiths-Jim describes.

144. In my opinion, the POSITA would have further appreciated that in order to time-align the microphone signals, for a sound that arrives at the first physical microphone before the second microphone, then a time *delay* would be applied to the *first microphone signal*, x_0 , to time-align it to the second microphone signal, x_1 . Moreover, as the POSITA would have understood, a delay would have been applied to the *first microphone signal* through shifting of the filter coefficients in the filter so that the filtered signal is also time-delayed.

145. Brandstein also expressly teaches modifying the original GSC to include additional delays to the microphone signals to compensate for errors in the assumed time alignment. For example, Brandstein explains that the basic GJBF suffers from target-signal cancellation due to steering-vector errors that can cause by an undesirable phase difference between the microphone signals. Ex. 1005, 90-91. As an example, Brandstein introduces delay elements z^{-L_2} for delaying microphone signals by L_2 samples. Ex. 1005, 92-93 & Fig. 5.4.

14. Claim 12: “The microphone array of claim 11, wherein the delay is raised to a power that is proportional to a time difference between arrival of the speech at the first virtual microphone and arrival of the speech at the second virtual microphone.”

146. In my opinion, Brandstein discloses or renders obvious this limitation.

147. The '691 patent defines a value γ as follows:

$$\gamma = \frac{d_2 - d_1}{c} f_s \text{ (samples),}$$

where “[t]he distances d_1 and d_2 are the distance from O_1 and O_2 to the speech source . . . , and γ is their difference divided by c , the speed of sound, and multiplied by the sampling frequency f_s .” Ex. 1001, 10:50-11:4. The '691 patent later refers to γ as “the time difference between arrival of speech at V_1 compared to V_2 .” Ex. 1001, 14:66-15:2.

148. It is my opinion that claim 12 would have been obvious to a POSITA based on Brandstein alone. A POSITA would have known that some applications involve near-field target signals, that the GSC requires time-alignment, that near-field time-alignment involves applying delays, and that time-aligning signals for near-field beamforming can be done by taking into account a standard near-field propagation model. The motivation and knowledge to use near-field time-alignment

to implement the delays for the GSC (as described in Brandstein's Chapter 5) would have been immediately apparent to a POSITA.

149. As Brandstein teaches, the GSC's BM has a delay-and-sum (or delay-and-subtract) structure because the microphone signals are delayed so that their target-signal components are aligned in time, which is what allows the subtraction operations in the BM to block the target signal. Ex. 1005, 88-89. Specifically, Brandstein states that "[d]elay-and-sum beamformers apply time shifts to the array signals to compensate for the propagation delays in the arrival of the source signal at each microphone. These signals are time-aligned and summed together to form a single output signal." Ex. 1005, 159. When the Griffiths-Jim paper published in 1982, delay-and-sum beamforming was already well known in the art, and POSITAs understood why Griffiths-Jim disclosed delay elements to time-align signals before feeding them to the BM. By 2007, it had been known for decades that time-alignment could be applied to the lower branch of the GSC.

150. In addition, in Chapter 5, Brandstein explains that the GSC's BM has a delay-and-subtract (delay-and-sum with negative coefficients) structure: "The BM in Fig. 5.1 has a simple delay-and-sum structure" (Ex. 1005, 90); and the "simplest structure for the BM is a delay-and-subtract beamformer which was described in the

previous section” (Ex. 1005, 88). A POSITA would have understood that the “previous section” refers to Chapter 2, which is the last chapter of Brandstein before Chapter 5 to discuss delay-and-sum/subtract structures.

151. In section 2.4, Brandstein discloses “all designs presented in section 2.3 can be implemented by using the GSC-structure.” Ex. 1005, 35. For example, Brandstein teaches that a “multi-channel *time-aligned* input signal ... is projected onto the noise-only subspace^[footnote] by a blocking matrix **B**,” with the footnote stating: “Which means that the desired signal is spatially filtered out (blocked).” Ex. 1005, 34 & n.5 (emphasis added). A “well-known example” of **B** “is the original Griffith[s]-Jim matrix which subtracts two adjacent channels.” Ex. 1005, 34. Chapter 2 further states that there is a “close relationship between superdirective arrays and adaptive beamformers and their well-known implementation as a generalized sidelobe canceler” and that “nearfield assumptions ... can be easily included.” Ex. 1005, 36-37.

152. Furthermore, a POSITA would have known of the differences between far-field and near-field sound and how to time-align signals from near-field sources. For example, Brandstein Chapter 2 discloses that time-aligning signals for near-field beamforming takes into account a standard near-field propagation model:

- Nearfield design, including attenuation of the desired signal [14], [22]

$$\mathbf{d}^T = [a_0 \exp(-j\omega\tau_0), a_1 \exp(-j\omega\tau_1), \dots, a_{N-1} \exp(-j\omega\tau_{N-1})], \quad (2.28)$$

$$a_i = \frac{\|\mathbf{q} - \mathbf{p}_{ref}\|}{\|\mathbf{q} - \mathbf{p}_i\|}, \quad (2.29)$$

$$\tau_i = \frac{\|\mathbf{q} - \mathbf{p}_{ref}\| - \|\mathbf{q} - \mathbf{p}_i\|}{c}, \quad (2.30)$$

where $\|\mathbf{q} - \mathbf{p}_{ref}\|$ and $\|\mathbf{q} - \mathbf{p}_i\|$ denote the distance between the vector location of the source $\mathbf{q}$ and a reference sensor $\mathbf{p}_{ref}$, or the sensor $\mathbf{p}_i$, respectively.

Ex. 1005, 25. The microphone attenuation factor a_i in Brandstein equation (2.29) is a ratio of the distance between the speech source and a reference (*first*) microphone to the distance between the speech source and the i th (*e.g., second*) microphone. This attenuation factor is *multiplied* to the i th (*e.g., second*) microphone signal to achieve equal amplitudes across all channels.

153. Brandstein's Chapter 2 therefore would have provided a POSITA with express motivation to use Brandstein's disclosure of near-field time-alignment in section 2.3 to implement the delay for the GSC. Furthermore, in Chapter 13, Brandstein explains that delay-and-sum beamforming was "first developed on the basis of the far-field assumption, but may also be extended to near-field beamforming." Ex. 1005, 290 (internal citation omitted). In my opinion, a POSITA would have understood that Brandstein teaches applying the complex-exponential term specified in

Brandstein equation 2.28 to delay a signal in proportion to τ_i to time-align the signals by disclosing: (1) the delay-and-sum or delay-and-subtract structure for the FBF and BM, respectively; (2) the purpose of the delay in delay-and-sum beamforming is to time-align the sensor signals; (3) delay-and-sum beamforming can be extended to near-field designs; and (4) the standard near-field design specifies a delay whose magnitude is equal to the difference in propagation time of the target signal to one microphone relative to another (reference) microphone. Combining the teachings of Brandstein Chapters 2 and 5 to arrive at claim 12 is a simple addition of one known element (near-field time alignment) to another known element (the GSC) to obtain predictable results (*e.g.*, effective target-signal blocking in the lower branch). The combination also applies a known technique of time-alignment for near-field sources to the well-known GSC that is ready for improvement, with the combination yielding predictable results. Thus, a POSITA would have had a reasonable expectation of success, and this limitation would have been obvious in view of Brandstein's Chapters 2 and 5 and the background knowledge of a POSITA.

154. Brandstein renders obvious extending the standard GSC to a near-field design, including by time-aligning the microphone signals fed to the BM by applying delay term $e^{-j\omega\tau_i}$, whose exponent (*power*) is *proportional* to τ_i . Brandstein defines τ_i as:

$$\tau_i = \frac{\|q - p_{ref}\| - \|q - p_i\|}{c}, \quad (2.30)$$

where $\|q - p_{ref}\|$ and $\|q - p_i\|$ denote the distance between the vector location of the source q and a reference sensor p_{ref} , or the sensor p_i , respectively.

Ex. 1005, 25.

155. Brandstein's τ_i results from dividing the *difference* of the distance between the source (i.e. *speech*) q and a reference sensor (*first virtual microphone*) and the distance between the source and the i th sensor (*second virtual microphone*) by the speed of sound. Brandstein's τ_i is therefore *a time difference between arrival of the speech at the first virtual microphone and arrival of the speech at the second virtual microphone*.

15. Claim 13: “The microphone array of claim 11, wherein the delay is raised to a power that is proportional to a sampling frequency multiplied by a quantity equal to a third distance subtracted from a fourth distance, the third distance being between the first physical microphone and the speech source and the fourth distance being between the second physical microphone and the speech source.”

156. In my opinion, Brandstein discloses or renders obvious this limitation.

157. As explained above with respect to claim 12, Brandstein's τ_i results from subtracting the distance between the source (i.e. *speech*) q and the i th sensor (*first physical microphone*), which is the *third distance*, from the distance between

the source and a reference sensor (*second physical microphone*), which is the *fourth distance*, and then dividing by the speed of sound, c . Therefore, τ_i is proportional to a quantity equal to a third distance subtracted from a fourth distance. Ex. 1005, Eqs. 2.28-2.30

158. Claim 13 further recites an exponent “proportional” to “a sampling frequency multiplied by” a quantity such as Brandstein’s τ_i . But τ_i is proportional to the sampling frequency multiplied by τ_i , so Brandstein’s delay term $e^{-j\omega\tau_i}$ satisfies claim 13.

- Nearfield design, including attenuation of the desired signal [14], [22]

$$\mathbf{d}^T = [a_0 \exp(-j\omega\tau_0), a_1 \exp(-j\omega\tau_1), \dots, a_{N-1} \exp(-j\omega\tau_{N-1})], \quad (2.28)$$

$$a_i = \frac{\|\mathbf{q} - \mathbf{p}_{ref}\|}{\|\mathbf{q} - \mathbf{p}_i\|}, \quad (2.29)$$

$$\tau_i = \frac{\|\mathbf{q} - \mathbf{p}_{ref}\| - \|\mathbf{q} - \mathbf{p}_i\|}{c}, \quad (2.30)$$

where $\|\mathbf{q} - \mathbf{p}_{ref}\|$ and $\|\mathbf{q} - \mathbf{p}_i\|$ denote the distance between the vector location of the source $\mathbf{q}$ and a reference sensor $\mathbf{p}_{ref}$, or the sensor $\mathbf{p}_i$, respectively.

Ex. 1005, 25.

159. Additionally, although Brandstein discloses the delay term in the form above with time expressed in seconds, a POSITA would have known, as a matter of

routine practice, how to rewrite the formula to express time in terms of number of samples, as is common for digital signal processing, which involves discrete sampling. To convert from seconds to samples, a POSITA would simply multiply the delay expressed in seconds by the sampling frequency f_s expressed in samples per second. For example, after that trivial operation, Brandstein's delay term $e^{-j\omega\tau_i}$ becomes $e^{-j\omega\tau_i f_s}$. Using z-transform notation, this is commonly written as $z^{-\tau_i f_s}$. Applying that exponential term delays a signal by $\tau_i f_s$ samples, as desired.

160. In my opinion, a POSITA would have known to rewrite Brandstein's near-field delay in terms of samples. Brandstein itself writes this delay element in terms of samples for use with the GSC because Fig. 5.1 is shown in terms of samples. Ex. 1005, 88-89 & n.1. Brandstein explains the effect of z^{-L_1} block in Fig. 5.1 as an " L_1 -sample delay". Brandstein also discloses implementing the GSC with a digital signal processor. Ex. 1005, 104-105, Fig. 5.16. A POSITA would have been motivated to implement the GSC digitally because digital signal processing was a known technique, and, among other advantages over analog systems, digital circuits can be reproduced more easily in large quantities at low cost, digital systems are unaffected by changes of component values in analog systems caused by factors such as age, temperature, and humidity, digital signal processing allows for flexibility through digitally programmable systems, digital systems are generally more accurate, digital

signals can be easily stored on media using semiconductor chips, and digital signals can be transmitted over long distances. Converting Brandstein's delay formula from seconds to samples for digital signal processing would have involved the simple and well-known operation of multiplying by the sampling frequency to convert time units from, e.g., seconds to samples. Finally, a POSITA would have understood that the labeling of Brandstein's *ith* and reference sensors as the *first and second microphones* (respectively) could be flipped, without affecting the analysis because (1) all of the requirements of the claims from which claim 13 depends (claims 1, 3, 4, 8, 9, 10, 11) with respect to both *microphones* are identical and apply to both *microphones* symmetrically and (2) the mapping for the limitation of claim 13 to Brandstein's Eqs. 2.28-2.30 applies to either *microphone*. Specifically, regarding (2), the POSITA would have readily appreciated that Brandstein's Eqs. 2.28-2.30 may be rewritten so that τ_i is proportional to the distance between q and the *ith* sensor, minus the distance between q and the reference sensor. This would have meant that the exponents in Eq. 2.28 would be rewritten to remove the negative sign in the exponential. In this case, the distance between q and the reference sensor is the *third distance*, and the distance between q and the *ith* sensor is the *fourth distance*.

16. **Claim 14: “The microphone array of claim 10, wherein the second microphone signal is multiplied by a ratio, wherein the ratio is a ratio of a third distance to a fourth distance, the third distance being between the first physical microphone and the speech source and the fourth distance being between the second physical microphone and the speech source.”**

161. Claim 14 requires the second microphone signal to be multiplied by a ratio of the distance between the speech source and the first physical microphone to the distance between the speech source and the second physical microphone. In my opinion, Brandstein discloses this limitation.

162. As discussed above, the microphone attenuation factor a_i in Brandstein equation (2.29):

- Nearfield design, including attenuation of the desired signal [14], [22]

$$\mathbf{d}^T = [a_0 \exp(-j\omega\tau_0), a_1 \exp(-j\omega\tau_1), \dots, a_{N-1} \exp(-j\omega\tau_{N-1})], \quad (2.28)$$

$$a_i = \frac{\|\mathbf{q} - \mathbf{p}_{ref}\|}{\|\mathbf{q} - \mathbf{p}_i\|}, \quad (2.29)$$

$$\tau_i = \frac{\|\mathbf{q} - \mathbf{p}_{ref}\| - \|\mathbf{q} - \mathbf{p}_i\|}{c}, \quad (2.30)$$

where $\|\mathbf{q} - \mathbf{p}_{ref}\|$ and $\|\mathbf{q} - \mathbf{p}_i\|$ denote the distance between the vector location of the source $\mathbf{q}$ and a reference sensor $\mathbf{p}_{ref}$, or the sensor $\mathbf{p}_i$, respectively.

Ex. 1005, 25. The microphone attenuation factor a_i in Brandstein equation (2.29) is a ratio of the distance between the speech source and a reference (*first*) microphone

to the distance between the speech source and the i th (e.g., *second*) microphone. This attenuation factor is *multiplied* to the i th (e.g., *second*) microphone signal to achieve equal amplitudes across all channels.

17. Claim 15: “The microphone array of claim 9, wherein the second virtual microphone comprises the first microphone signal subtracted from the second microphone signal.”

163. In my opinion, Brandstein discloses or renders obvious this limitation.

164. The POSITA would have understood that the labeling of Brandstein’s microphone signals x_0 and x_1 as the *first and second microphone signals* (respectively) could be flipped, without affecting the analysis. This is because all of the requirements in claim 15 (including claims 1, 3, 4, 8, 9, from which claim 15 depends) with respect to both *microphone signals* are identical and apply to both *microphone signals* symmetrically

165. For this claim, microphone signal x_1 is the “first” microphone signal and microphone signal x_0 is the “second.” With this designation, the *second virtual microphone* includes the BM with its delay-and-subtract structure that *subtracts* the *first microphone signal* from the *second microphone signal*:

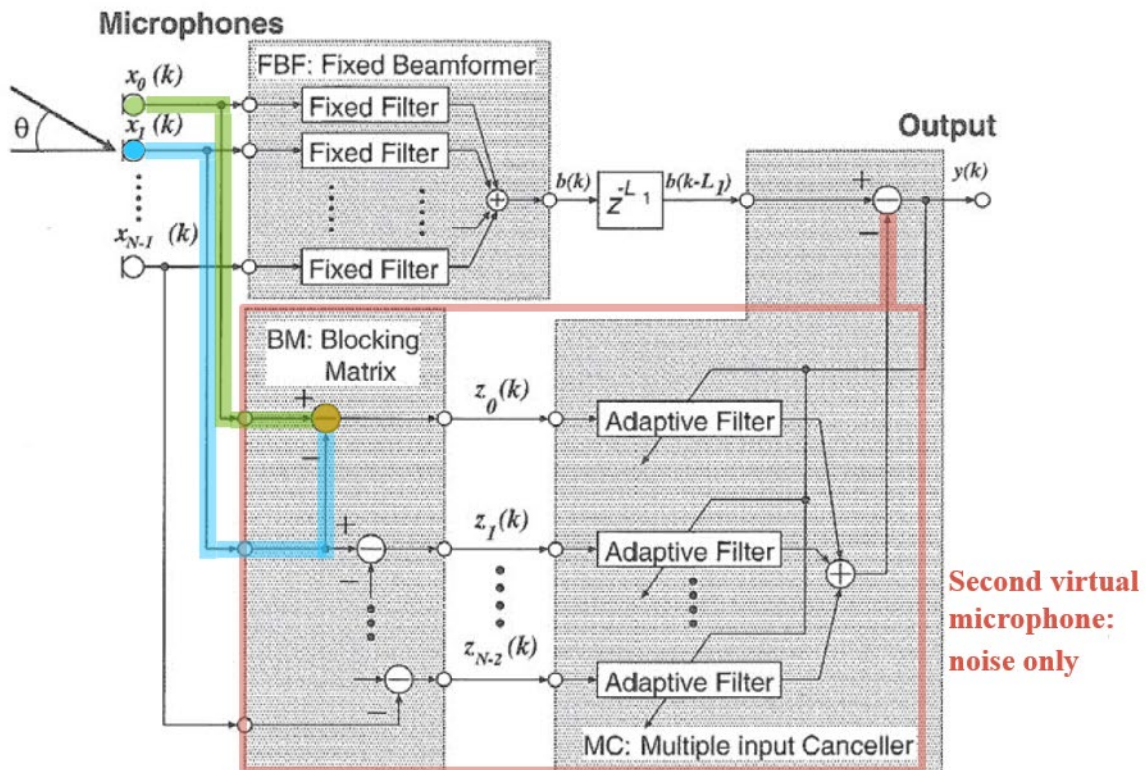


Fig. 5.1. Griffiths-Jim beamformer. It comprises a fixed beamformer (FBF), a multiple-input canceler (MC), and a blocking matrix (BM).

Ex. 1005, 89, Fig. 5.1 (annotated)

18. Claim 16: “The microphone array of claim 15, wherein the first microphone signal is delayed.”

166. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claims 11 and 15.

19. **Claim 17: “The microphone array of claim 16, wherein the delay is raised to a power that is proportional to a time difference between arrival of the speech at the first virtual microphone and arrival of the speech at the second virtual microphone.”**

167. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 12.

20. **Claim 18: “The microphone array of claim 16, wherein the power is proportional to a sampling frequency multiplied by a quantity equal to a third distance subtracted from a fourth distance, the third distance being between the first physical microphone and the speech source and the fourth distance being between the second physical microphone and the speech source.”**

168. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 13.

21. **Claim 19: “The microphone array of claim 18, wherein the first microphone signal is multiplied by a ratio, wherein the ratio is a ratio of the third distance to the fourth distance.”**

169. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 14.

22. Claim 20: “The microphone array of claim 4, wherein the single null is located at a distance from at least one of the first physical microphone and the second physical microphone where the source of the speech is expected to be.”

170. In my opinion, Brandstein discloses or renders obvious this limitation.

171. Brandstein’s Fig. 5.3 bottom directivity plot shows that the single null is in the direction of the target signal (*speech*). This null occurs at all distances in the direction of the speech, and therefore is located at *a distance from at least one of the first physical microphone and the second physical microphone where the source of the speech is expected to be*.

23. Claim 21: “The microphone array of claim 1, wherein the first virtual microphone comprises the second microphone signal subtracted from a delayed version of the first microphone signal.”

172. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claims 10 and 11.

173. For the same reasons as explained above for claim 10, the *first microphone signal* and the *second microphone signal* are input into filters in the top branch (*first virtual microphone*) of the GSC and then summed. A POSITA would understand that the filters apply weight coefficients to the incoming microphone signals,

and that these weight coefficients typically include both positive and negative weights.

174. Applying a negative weight to the *second microphone signal* would *subtract* it from the *first microphone signal*. Furthermore, as explained above, the filters are implemented by multiplying the weights times *delayed* versions of each *microphone signal*. Thus, as explained above, the output of the FBF, “ $b(k)$ ”, *comprises the second microphone signal* (“ x_1 ”) *subtracted from a delayed version of the first microphone signal* (“ x_0 ”).

24. Claim 22: “The microphone array of claim 21, wherein the second virtual microphone comprises a delayed version of the first microphone signal subtracted from the second microphone signal.”

175. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claims 10, 11, and 15

176. As explained for claim 15, the *second virtual microphone comprises the first microphone signal subtracted from the second microphone signal*, and as explained for claim 11 the *first microphone signal is delayed*. The POSITA would have understood that the labeling of Brandstein’s microphone signals x_0 and x_1 as the *first and second microphone signals* (respectively) could be flipped, without affecting the analysis because all of the requirements in claim 22 (including claims 1

and 21 from which claim 22 depends) with respect to both *microphone signals* are identical and apply to both *microphone signals* symmetrically. Thus, by labeling x_1 as the *first microphone signal* and x_0 as the *second microphone signal*, the BM in the *second virtual microphone subtracts* the *first microphone signal* from the *second microphone signal*, and the *first microphone signal* is *delayed* for time-alignment.

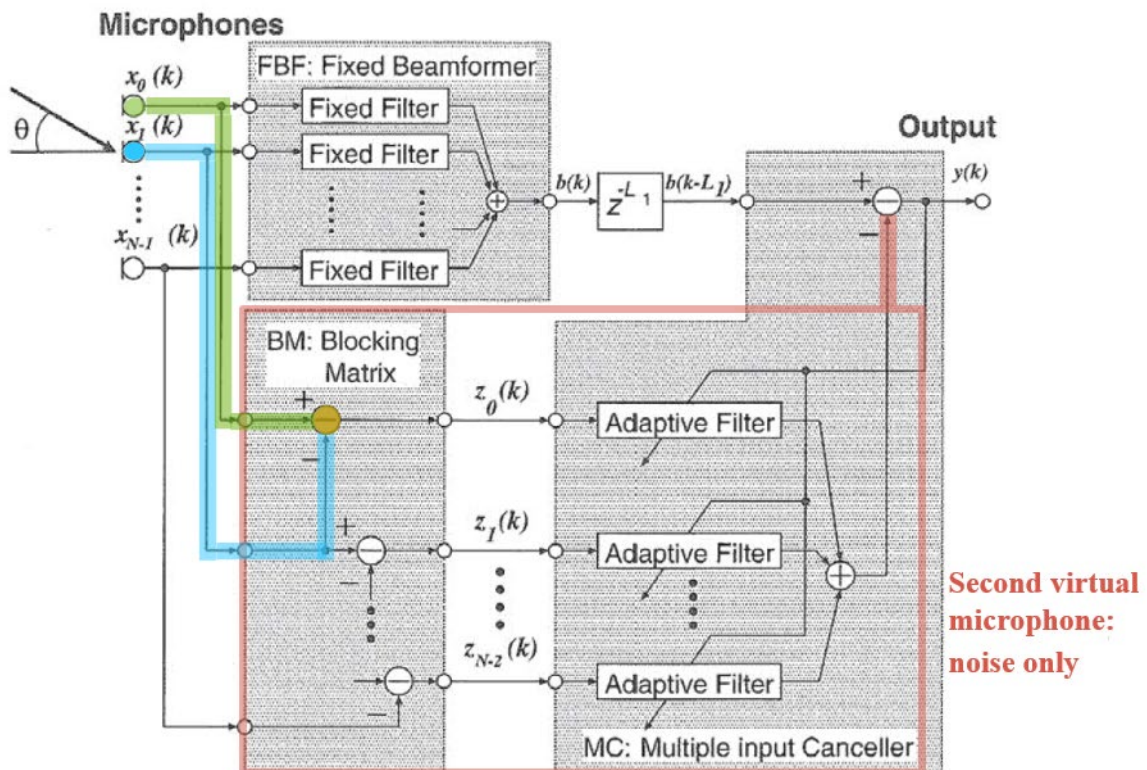


Fig. 5.1. Griffiths-Jim beamformer. It comprises a fixed beamformer (FBF), a multiple-input canceller (MC), and a blocking matrix (BM).

Ex. 1005, 89, Fig. 5.1 (annotated)

25. Claim 23

a) **[23pre]: “A microphone array comprising:”**

177. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1pre].

b) **[23a]: “a first virtual microphone formed from a first combination of a first microphone signal and a second microphone signal,”**

178. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1a].

c) **[23b]: “wherein the first microphone signal is generated by a first omnidirectional microphone and the second microphone signal is generated by a second omnidirectional microphone; and”**

179. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claims [1d] and 2.

d) **[23c]: “a second virtual microphone formed from a second combination of the first microphone signal and the second microphone signal,”**

180. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1e].

- e) **[23d]: “wherein the second combination is different from the first combination;”**

181. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1e].

- f) **[23e]: “wherein the first virtual microphone has a first linear response to speech that is substantially similar across a plurality of frequencies for a speech source within a predetermined angle relative to an axis of the microphone array and devoid of a null, and a first linear response to noise that is devoid of a null,”**

182. In my opinion, Brandstein renders obvious this limitation for the same reasons discussed above regarding claims [1c] and 3.

183. Brandstein Fig. 5.10 is a plot the modified-GJBF’s sensitivity showing that the sensitivity is constant within an angle of $\theta = -20^\circ$ to 20° , where θ is *an angle relative to an axis of the microphone array* as shown in Fig. 5.1. For the simulation that resulted in Fig. 5.10, Brandstein states that the “assumed target direction [i.e. *speech source*] was 0° ” and that the “maximum allowable target-direction error” was 20° . Ex. 1005, 100. In my opinion, a POSITA would have understood the “maximum allowable target-direction error” refers to a pre-set threshold, which means it is *a predetermined angle* defining a range of angles *within which a speech source is located*.

184. Furthermore, as can be seen in the directivity pattern of the FBF in Fig. 5.3, there is no zero or minima in the spatial response in the direction of interference (noise), i.e. *first linear response to noise*. Therefore, the basic GSC has a *first linear response to noise* that is devoid of a null.

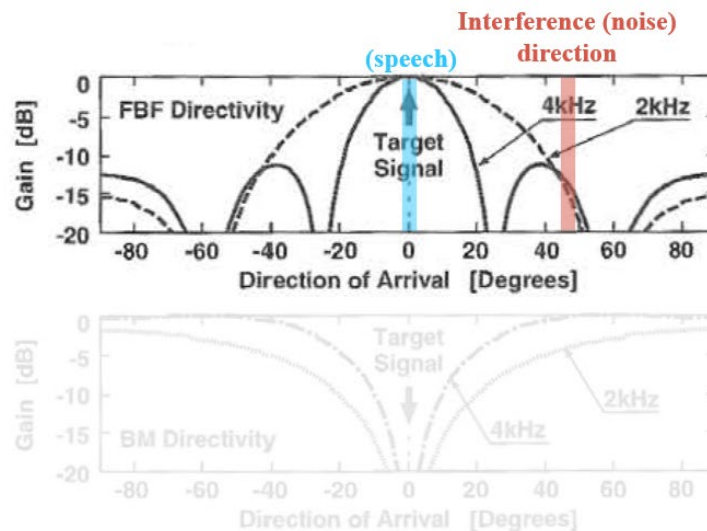


Fig. 5.3. Directivity pattern of a fixed beamformer (FBF) and a blocking matrix (BM).

Ex. 1005, Fig. 5.3 (annotated)

- g) [23f]: “wherein the second virtual microphone has a second linear response to speech that has a single null oriented in a direction toward a source of the speech and a second linear response to noise that is devoid of a null,”

185. In my opinion, Brandstein discloses or renders obvious this limitation.

186. Brandstein discloses or renders obvious *a second linear response to speech that has a single null oriented in a direction toward a source of the speech* for the same reasons discussed above regarding claim 4.

187. Furthermore, as can be seen in the directivity pattern of the BM (part of the *second virtual microphone*), there is no zero or minima in the spatial response in the direction of interference (*noise*), i.e. *second linear response to noise*. Therefore, for the same reasons discussed above regarding claim 3, the simple GJBF has a *second linear response to noise* that is devoid of a null.

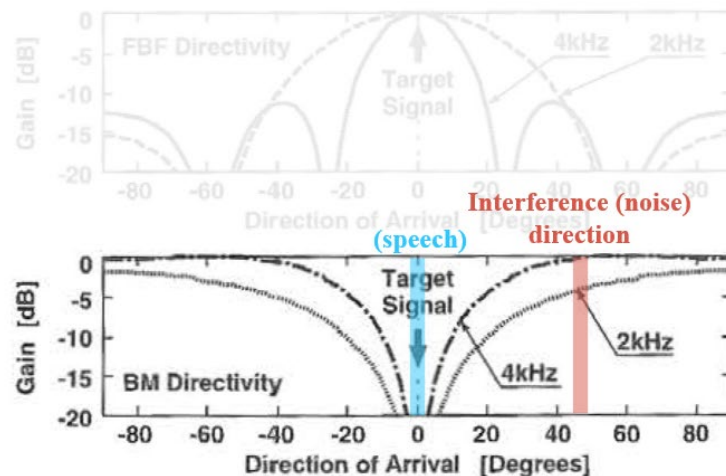


Fig. 5.3. Directivity pattern of a fixed beamformer (FBF) and a blocking matrix (BM).

Ex. 1005, Fig. 5.3 (annotated)

- h) **[23g]: “wherein the second linear response to noise is substantially similar to the first linear response to noise**

and the second linear response to speech is substantially dissimilar to the first linear response to speech,”

188. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1g].

i) **[23h]: “wherein the speech is human speech.”**

189. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 3.

26. Claim 24: “The microphone array of claim 23, wherein the single null is a region of the second linear response to speech having a measured response level that is lower than the measured response level of any other region of the second linear response to speech.”

190. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 5.

27. Claim 25: “The microphone array of claim 23, wherein the second linear response to speech includes a primary lobe oriented in a direction away from the source of the speech.”

191. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 6.

- 28. Claim 26: “The microphone array of claim 25, wherein the primary lobe is a region of the second linear response to speech having a measured response level that is greater than the measured response level of any other region of the second linear response to speech.”**

192. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 7.

29. Claim 27

- a) **[27pre]: “A device comprising:”**

193. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1pre]. In particular, a POSITA would have known that the GJBF and modified GJBF are both *devices*.

- b) **[27a]: “a first microphone outputting a first microphone signal and a second microphone outputting a second microphone signal; and”**

194. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1d]. A POSITA have understood that the first microphone signal is generated as the first microphone’s output, and the second microphone signal is generated as the second microphone’s output.

- c) **[27b]: “a processing component coupled to the first microphone signal and the second microphone signal,”**

195. In my opinion, Brandstein discloses or renders obvious this limitation.

196. As shown in Fig. 5.1, the top and bottom branches, including the FBF and BM, are *processing components* that are *coupled to the first and second microphone signals*. The components of both branches take as inputs the first and second microphone signals, and processes those signals through filtering and adding, among other things, to form virtual microphone signals.

- d) **[27c]: “the processing component generating a virtual microphone array comprising a first virtual microphone and a second virtual microphone,”**

197. In my opinion, Brandstein discloses or renders obvious this limitation.

198. The top and bottom branches of the GSC *generate virtual microphone* signals, which as discussed for claim [1b] and [1e], are *a first virtual microphone and a second virtual microphone*. These first and second microphones form an *array*.

- e) **[27d]: “wherein the first virtual microphone comprises a first combination of the first microphone signal and the second microphone signal,”**

199. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1a].

- f) **[27e]: “wherein the second virtual microphone comprises a second combination of the first microphone signal and the second microphone signal,”**

200. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1e].

- g) **[27f]: “wherein the second combination is different from the first combination,”**

201. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1e].

- h) **[27g]: “the first virtual microphone having a first linear response to speech and a first linear response to noise, the first linear response to speech being substantially similar across a plurality of frequencies for a speech source at an angle relative to an axis of the microphone array,”**

202. In my opinion, Brandstein renders obvious this limitation for the same reasons discussed above regarding claims [1b] and [1c].

- i) **[27h]: “the second virtual microphone having a second linear response to speech and a second linear response to noise,”**

203. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1f].

- j) **[27i]: “the second linear response to noise being substantially similar to the first linear response to noise, and the second linear response to speech being**

substantially dissimilar to the first linear response to speech,”

204. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1g].

k) **[27j]: “wherein one of the first linear response to speech and the second linear response to speech is reduced in a direction towards a source of speech.”**

205. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [4]. As explained above, the second linear response to speech of the second virtual microphone includes a null oriented in a direction toward a source of the speech. This null is a *reduced* response *in a direction towards a source of speech*. Therefore, “one of the first linear response to speech and the second linear response to speech” is reduced in a direction towards a source of speech

30. Claim 28

a) **[28pre] “A device comprising:”**

206. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [27pre].

b) **[28a]: “a first microphone outputting a first microphone signal and a second microphone outputting a second microphone signal, wherein the**

**first microphone and the second microphone are
omnidirectional microphones; and”**

207. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claims [1d], [27a], and 2.

c) **[28b]: “a virtual microphone array comprising a first
virtual microphone and a second virtual microphone,”**

208. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claims [1b], [1e], and [27c].

d) **[28c]: “the first virtual microphone having a first
linear response to speech based on distances from a
first physical microphone and a second physical
microphone to a source of speech, and**

209. In my opinion, Brandstein discloses or renders obvious this limitation. As I discuss above for [1b], Brandstein teaches a *first virtual microphone having a first linear response to speech... a first linear response to noise*. With respect to the *first linear response to speech being based on distances from a first physical microphone and a second physical microphone to a source of speech*, the POSITA would have readily appreciated that a microphone’s *response to speech* is a function of the distance to the microphone, due to attenuation characteristics. The *first virtual microphone response to speech* is based on both of Brandstein’s *microphones’ distances* because both of Brandstein’s *microphones* (and therefore both *distances*) are

included in the *first virtual microphone*. Additionally, as I explain above for claims 12-13, the *first virtual microphone* applies time-alignment delays, where those delays are calculated based on the claimed *distances*, such that the *first virtual microphone* has a *first linear response to speech based on these distances*.

- e) **[28d] a first linear response to noise, the first linear response to speech being substantially similar across a plurality of frequencies for a speech source within a predetermined angle relative to an axis of the microphone array,”**

210. In my opinion, Brandstein renders obvious this limitation for the same reasons discussed above regarding claim [1c].

- f) **[28e]: “the second virtual microphone having a second linear response to speech and a second linear response to noise,”**

211. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1f].

- g) **[28f]: “the second linear response to noise being substantially the first linear response to noise, and the second linear response to speech being substantially dissimilar to the first linear response to speech,”**

212. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1g].

- h) **[28g]: “wherein the distances form a ratio with which to direct the first and second linear responses to speech toward the source of speech,”**

213. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claims 14 and 19. Brandstein discloses *the distances form a ratio* for the same reasons discussed above regarding claims 14 and 19. As explained in Section VII.A.13-15 (claims 11-13), this ratio is used to time-align the incoming microphone signals, which the POSITA would have understood is performed so that GJBF’s bottom branch can accurately null out the speech signal, while ensuring the GJBF’s top branch has a non-zero response toward the speech. Thus, the *ratio* in Brandstein is used to *direct the first and second linear responses to speech toward the source of speech*.

- i) **[28h]: “wherein the first virtual microphone comprises a first combination of the first microphone signal and the second microphone signal,”**

214. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1a].

- j) **[28i]: “wherein the second virtual microphone comprises a second combination of the first microphone signal and the second microphone signal,”**

215. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1e].

- k) **[28j]: “wherein the second combination is different from the first combination,”**

216. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1e].

- l) **[28k]: “wherein the first virtual microphone and the second virtual microphone are distinct virtual directional microphones.”**

217. In my opinion, Brandstein discloses this limitation.

218. The directivity patterns shown in Brandstein’s Fig. 5.3 shows that the responses of the FBF (of the *first virtual microphone*) and BM (of the *second virtual microphone*) are different (*distinct*) from each other and depend on the direction of arrival. Therefore, the FBF and BM have *directional* characteristics.

31. Claim 29

- a) **[29pre]: “A device comprising:”**

219. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [27pre].

- b) **[29a]: “a first physical microphone generating a first microphone signal;”**

220. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1d].

- c) **[29b]: “a second physical microphone generating a second microphone signal; and”**

221. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1d].

- d) **[29c]: “a processing component coupled to the first microphone signal and the second microphone signal,”**

222. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [27b].

- e) **[29d]: “the processing component generating a virtual microphone array comprising a first virtual microphone and a second virtual microphone,”**

223. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [27c].

- f) **[29e]: “the first virtual microphone having a first linear response to speech and a first linear response to noise,”**

224. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1b].

- g) **[29f]: “the first linear response to speech being substantially similar across a plurality of frequencies for a speech source at an angle relative to an axis of the microphone array,”**

225. In my opinion, Brandstein renders obvious this limitation for the same reasons discussed above regarding claim [1c].

- h) **[29g]: “the second virtual microphone having a second linear response to speech and a second linear response to noise,”**

226. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1f].

- i) **[29h]: “the second linear response to noise being substantially similar to the first linear response to noise, and the second linear response to speech being substantially dissimilar to the first linear response to speech,”**

227. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1g].

- j) **[29i]: “wherein at least one of the first linear response to noise and the second linear response to noise is non-zero in a direction towards a source of noise,”**

228. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1h].

- k) **[29j]: “wherein the first virtual microphone comprises the second microphone signal subtracted from a delayed version of the first microphone signal,”**

229. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claims, 10, 11, and 21.

- l) **[29k]: “wherein the second virtual microphone comprises a delayed version of the first microphone signal subtracted from the second microphone signal.”**

230. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 22.

- 32. Claim 30: “The device of claim 29, wherein the first linear response to speech is devoid of a null, wherein the speech is human speech.”**

231. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 3.

- 33. Claim 31: “The device of claim 30, wherein the second linear response to speech of the second virtual microphone includes a single null oriented in a direction toward a source of the speech.”**

232. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 4.

- 34. Claim 32: “The device of claim 31, wherein the single null is a region of the second linear response to speech having a measured response level that is lower than the measured response level of any other region of the second linear response to speech.”**

233. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 5.

- 35. Claim 33: “The device of claim 31, wherein the second linear response to speech includes a primary lobe oriented in a direction away from the source of the speech.”**

234. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 6.

- 36. Claim 34: “The device of claim 33, wherein the primary lobe is a region of the second linear response to speech having a measured response level that is greater than the measured response level of any other region of the second linear response to speech.”**

235. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 7.

- 37. Claim 35: “The device of claim 31, wherein the first physical microphone and the second physical microphone are positioned along an axis and separated by a first distance.”**

236. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 8.

- 38. Claim 36: “The device of claim 35, wherein a midpoint of the axis is a second distance from the speech source that generates the speech, wherein the speech source is located in a direction defined by an angle relative to the midpoint.”**

237. Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 9.

- 39. Claim 37: “The device of claim 36, wherein one or more of the first microphone signal and the second microphone signal is delayed.”**

238. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 11.

- 40. Claim 38: “The device of claim 37, wherein the delay is raised to a power that is proportional to a time difference between arrival of the speech at the first virtual microphone and arrival of the speech at the second virtual microphone.”**

239. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 12.

- 41. Claim 39: “The device of claim 38, wherein the power is proportional to a sampling frequency multiplied by a quantity equal to a third distance subtracted from a fourth distance, the third distance being between the first physical microphone and the speech source and the fourth distance being between the second physical microphone and the speech source.”**

240. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 18.

- 42. Claim 40: “The device of claim 36, wherein one or more of the first microphone signal and the second microphone signal is multiplied by a gain factor.”**

241. In my opinion, Brandstein renders obvious this limitation.

242. As I explained for claim 14, Brandstein's microphone attenuation factor a_i is *multiplied* by the *i*th (e.g., second) microphone signal to obtain equal amplitudes across all channels.

243. Furthermore, as the original Griffiths-Jim paper discusses, the sensor outputs were combined by multiplying the sensor output signals by factors called "weights" ($w_{c1}, \dots, w_{cM}$ in the paper), which are also called "*gains*" (or *gain factors*) and further filtering and summing the weighted sensor signals. As discussed in the Background Section, Brandstein applies the teachings of Griffiths-Jim to the microphone context.

43. Claim 41

a) **[41pre]: "A sensor comprising:"**

244. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1pre]. A POSITA would have understood that the microphone array and the associated signal processing for combining the output signals constitute a sensor.

b) **[41a]: "a physical microphone array including a first physical microphone and a second physical microphone,"**

245. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claims [1pre] and [1d].

- c) **[41b]: “the first physical microphone outputting a first microphone signal and the second physical microphone outputting a second microphone signal; and”**

246. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claims [1d], [27a], and 2.

- d) **[41c]: “a virtual microphone array comprising a first virtual microphone and a second virtual microphone,”**

247. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [27c].

- e) **[41d]: “the first virtual microphone comprising a first combination of the first microphone signal and the second microphone signal and”**

248. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1a].

- f) **[41e]: “having a first linear response to speech and a first linear response to noise, the first linear response to speech being substantially similar across a plurality of frequencies for a source of speech at an angle relative to an axis of the microphone array,”**

249. In my opinion, Brandstein renders obvious this limitation for the same reasons discussed above regarding claims [1b] and [1c].

- g) **[41f]: “the second virtual microphone comprising a second combination of the first microphone signal and the second microphone signal and”**

250. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1e].

- h) **[41g]: “having a second linear response to speech and a second linear response to noise,”**

251. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1f].

- i) **[41h]: “the second linear response to noise being substantially similar to the first linear response to noise and the second linear response to speech being substantially dissimilar to the first linear response to speech,”**

252. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [1g].

- j) **[41i]: “wherein one of the first linear response to speech and the second linear response to speech is reduced in a direction towards a source of speech,”**

253. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim [27j].

- k) **[41j]: “wherein the virtual microphone array includes a single null oriented in a direction toward the source of speech of a human speaker.”**

254. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claims 4, [23e]-[23f], and 3.

44. **Claim 42: “The sensor of claim 41, wherein the first linear response to speech is devoid of a null, wherein the second linear response to speech of the second virtual microphone includes the single null.”**

255. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claims 3 and 4.

45. **Claim 43: “The sensor of claim 42, wherein the single null is a region of the second linear response to speech having a measured response level that is lower than the measured response level of any other region of the second linear response to speech.”**

256. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 5.

46. **Claim 44: “The sensor of claim 42, wherein the second linear response to speech includes a primary lobe oriented in a direction away from the source of the speech.”**

257. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 6.

47. Claim 45: “The sensor of claim 44, wherein the primary lobe is a region of the second linear response to speech having a measured response level that is greater than the measured response level of any other region of the second linear response to speech.”

258. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 7.

48. Claim 46: “The sensor of claim 41, wherein the single null is located at a distance from the physical microphone array where the source of the speech is expected to be.”

259. In my opinion, Brandstein discloses or renders obvious this limitation for the same reasons discussed above regarding claim 20.

B. Ground 2: Brandstein in combination with McCowan

260. For the reasons I explain below, it is my opinion that claims 12-14, 17-19, 28, and 38-39 would have been obvious in view of Brandstein and McCowan.

1. Motivation to Combine

261. Applying near-field time-alignment for the lower branch of the GSC would have been part of a POSITA’s background knowledge because McCowan discloses this:

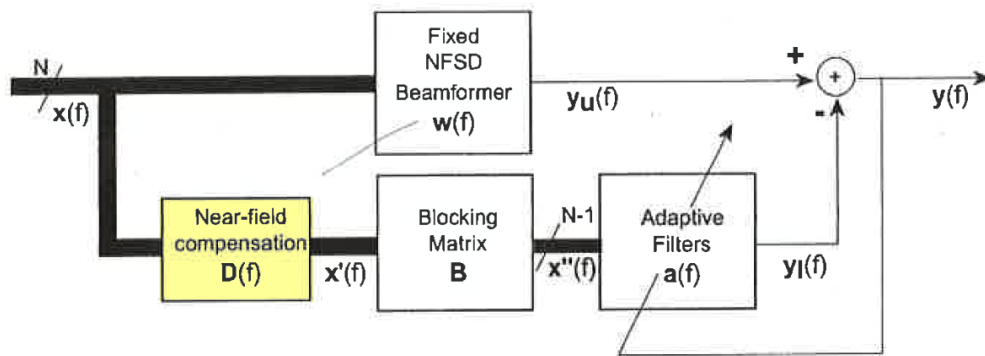


FIG. 3. Near-field adaptive beamformer.

Ex. 1007, 91, Fig. 3.

262. McCowan discloses an adaptive noise cancellation system for near-field speech sources. Ex. 1007, 87-89, 97. Fig. 3 shows that the input $x(f)$ including signals from multiple microphones is processed via two signal paths. Ex. 1007, 90, 93 (equation 15), 95-96. The BM “block[s] the desired signal from the adaptive noise estimate” by subtracting one microphone signal from another. Ex. 1007, 93, 95 (equations 18, 19). As McCowan explains, “[t]o ensure complete blocking, the desired signal must both be time aligned and have equal amplitudes across all channels.” (Ex. 1007, 93.)

263. McCowan teaches that for a near-field sound source, phase misalignment across microphones results from “[t]he difference in propagation time to [the i th] microphone with respect to the reference microphone ($i=1$),” defined as:

$$\tau_i = \frac{d_i - d_1}{c}, \quad (3)$$

where d_i is the distance between the source and the i th microphone and c is the speed of sound. Ex. 1007, 89-90. Additionally, amplitude differences across microphones are “negligible for far-field sources, but can be significant in the near-field case.” Ex. 1007, 90. The attenuation factor of wave amplitude on the i th microphone, with respect to the amplitude on the reference microphone ($i=1$), is defined as:

$$\alpha_i = \frac{d_1}{d_i}. \quad (4)$$

Ex. 1007 90.

264. McCowan explains that if $x_1(f)$ is defined as the desired source at a reference microphone, then the signal on the i th microphone is given by the equation:

$$x_i(f) = \alpha_i x_1(f) e^{-j2\pi f \tau_i}. \quad (5)$$

Ex. 1007 90.

265. Fig. 3 of McCowan shows that “[o]nce this near-field compensation has been performed, a standard GSC blocking matrix can be employed to block the desired signal from the adaptive path.” Ex. 1007, 93. McCowan also states that “[t]he blocking matrix and adaptive noise canceling filters are taken from the standard GSC technique [citation]” Ex. 1007, 95 (citing Griffiths-Jim).

266. In view of this background knowledge, it is my opinion that claims 12-14, 17-19, 28, and 38-39 would have been obvious to a POSITA based on Brandstein. A POSITA would have known that some applications involve near-field target signals, that the GSC’s BM requires time-alignment, that near-field time-alignment involves applying delays, and that McCowan discloses such near-field time-alignments followed by a standard GSC BM. The motivation and knowledge to use near-field time-alignment to implement the delays for the delay-and-subtract BM in the original Griffiths-Jim GSC (as described in Brandstein’s Chapter 5) would have been immediately apparent to a POSITA.

267. Alternatively, to the extent near-field time-alignment would require modifying the GSC described in Brandstein’s Chapter 5, such modification represents nothing more than the simple addition of one known element (near-field time alignment) to another known element (the GSC) to obtain the predictable results of

effective target-signal blocking in the lower branch (to permit effective noise cancellation in the final output without canceling the target signal).

2. Expectation of Success

268. In my opinion, the combination of McCowan and Brandstein applies a known technique (time-alignment for near-field sources) to a well-known device (GSC) that is ready for improvement, with the combination yielding predictable results. A POSITA would have had a reasonable expectation of success because, as McCowan demonstrates, near-field time-alignment with the GSC was trivial and shown to work. Claim 13 would therefore have been obvious in view of Brandstein's Chapter 5 and the background knowledge of a POSITA.

3. Claim 12

269. In my opinion, Brandstein in combination with McCowan renders this limitation obvious for the reasons discussed above with respect to Ground 1 claim 12 and for the additional reasons below.

270. McCowan defines τ_i as:

$$\tau_i = \frac{d_i - d_1}{c} \quad (3)$$

Ex. 1007, 89. In the above equations, c is the speed of sound. For McCowan, d_1 is the distance from the *speech source* to the reference (i.e. *first*) microphone and d_i is the distance from the speech source to the i th (e.g., *second*) microphone. The difference between d_i and d_1 is divided by the speed of sound to obtain *a time difference between arrival of the speech at the first virtual microphone and arrival of the speech at the second virtual microphone*, τ_i .

4. Claim 13

271. In my opinion, Brandstein in combination with McCowan renders this limitation obvious for the reasons discussed above with respect to Ground 1 claim 13 and for the additional reasons below.

272. Claim 13 recites an exponent “proportional” to “a sampling frequency multiplied by” a quantity such as Brandstein’s and McCowan’s τ_i . But τ_i is proportional to the sampling frequency multiplied by τ_i , so McCowan’s delay term $e^{-j\omega\tau_i}$ satisfies claim 13.

273. Alternatively, while McCowan discloses the delay term in the form above with time expressed in seconds, for the same reasons as discussed above with respect to Ground 1 claim 13, it would have also been routine for a POSITA to re-write the formula to express time in terms of number of samples, as is common for digital signal processing, which involves discrete sampling.

5. Claim 14

274. In my opinion, Brandstein in combination with McCowan renders this limitation obvious for the reasons discussed above with respect to Ground 1 claim 14 and for the additional reasons below.

275. McCowan equation (4) gives the microphone attenuation factor a_i :

$$\alpha_i = \frac{d_1}{d_i}, \quad (4)$$

is a ratio of d_1 (the distance between the speech source and a reference (*first*) microphone) to d_i (the distance between the speech source and the i th (*e.g., second*) microphone). Ex. 1009, 90. This attenuation factor is *multiplied* to the i th (*e.g., second*) microphone signal to achieve equal amplitudes across all channels.

6. Claim 17

276. In my opinion, Brandstein in combination with McCowan renders this limitation obvious for the reasons discussed above with respect to Ground 1 claim 17 and Ground 2 claim 12.

7. Claim 18

277. In my opinion, Brandstein in combination with McCowan renders this limitation obvious for the reasons discussed above with respect to Ground 1 claim 18 and Ground 2 claim 13.

8. Claim 19

278. In my opinion, Brandstein in combination with McCowan renders this limitation obvious for the reasons discussed above with respect to Ground 1 claim 19 and Ground 2 claim 14.

9. Claim 28

279. In my opinion, Brandstein in combination with McCowan renders this limitation obvious for the reasons discussed above with respect to Ground 1 claim 28 and Ground 2 claims 14 and 19.

10. Claim 38

280. In my opinion, Brandstein in combination with McCowan renders this limitation obvious for the reasons discussed above with respect to Ground 1 claim 28 and Ground 2 claim 12.

11. Claim 39

281. In my opinion, Brandstein in combination with McCowan renders this limitation obvious for the reasons discussed above with respect to Ground 1 claim 39 and Ground 2 claim 13.

IX. CONCLUSION

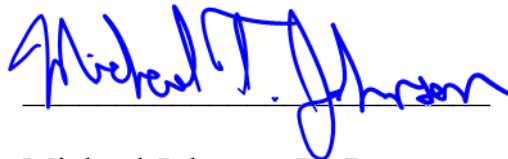
282. For the foregoing reasons, it is my opinion that claims 1-46 of the ‘691 patent would have been obvious to a person of ordinary skill in the art at the time of the alleged invention in view of the prior art discussed above.

283. I reserve the right to supplement my opinions in the future to address or respond to any issues that the Patent Owner may raise, as well as new information including, but not limited to, any claim constructions advanced by the Patent Owner or adopted by the Board in the Institution Decision, and to respond to any alleged secondary considerations as they become available to me.

284. I declare that all statements made herein are of my own knowledge and are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Declaration of Dr. Michael Johnson in Support of
Petition for *Inter Partes* Review of U.S. Patent No. 8,503,691

Executed on December 20, 2023.



Michael Johnson, Ph.D.