# A METHOD FOR MOVING VIRTUAL MICROPHONES USING ACOUSTIC MEASUREMENT DATA FROM MUSICAL INSTRUMENTS

D Carugo Oxford Brookes University, Oxford, UK.

## 1 INTRODUCTION

Research on the acoustic radiation characteristics of musical instruments goes back decades, for example in the work of Wogram and Meyer on wind instruments[1, 2], which itself references earlier work by Meyer from 1965. This paper contributes to the body of work by outlining a method that can be implemented in software for using the 3D acoustic radiation measurement data from musical instruments (generalized into instrument types) to process recorded sounds from that instrument type with the resulting signal having timbral characteristics as if the recording microphone were originally placed at a point that the operator of the software chooses.

Although instrument radiation data exists for a number of instrument types, for example in the work of Cook and Trueman[3], Pätynen and Lokki[4], and Pelzer et al[5], the spatial density of the data from these measurements varies by measurement apparatus and experimental approach. While it is possible to interpolate data between measurement points, some data may be lost in the high frequency range[6] when doing so. There is a minimum spatial density of measurement points (in terms of radial angle) that is needed in order to adequately capture data from the high frequency audio range which may radiate in narrow lobes.

Some researchers such Campbell[7] have used mechanical stimulation of the instrument to measure acoustic radiation patterns and thus derive directivity information for different frequencies, as the human musician is considered to be unable to make consistent soundings of the instrument. This allows the measurement apparatus used to take directional measurements individually, as the instrument stimulus is consistent. Other researchers (e.g. Behler et al[8]) have used the musician to sound the instrument as they may influence the radiation pattern with their presence. When the musician is used to sound the instrument for measurement, the general approach has been to use a microphone array to surround the musician/instrument and to take simultaneous measurements which can be compared, as it is not thought that the musician can accurately repeat their sounding of the instrument to the degree required for valid comparative measurements. The JASA article by Shabtai et al[9] contains a comparison of the violin radiation patterns with and without the player in position, so that the effect of the musician can be observed.

Some applications of 3D acoustic radiation data include sound synthesis: using a 3D capable loudspeaker system would allow the synthesized instrument to excite the listening space as a real instrument would in that same space[3, 10]. Ismail et al describe a similar application to enhance the Vsound violin model described in their 2017 AES article[11]. Another application used by researchers for this kind of data set is to more accurately model concert hall acoustics using virtual instruments within the model as described by Pedersini et al in 2000[12] and by Pätynen and Lokki in 2016[13].

A farther application for this type of musical instrument acoustic radiation data set is described in this paper which uses the data set for an instrument type (the example described in this paper is an acoustic guitar) to derive a transfer function that can be applied to modify a recorded sound so that it seems as if the recording microphone had been placed in another position in relation to the instrument.

# 2 MUSICAL INSTRUMENT RADIATION DATA

## 2.1 Data sets and current applications

Over time, a number of different sets of musical instrument radiation data have been produced by a variety of researchers. In addition to the early work by Meyer[14], research at Princeton by Cook and Trueman[3] resulted in a data set of various instruments, and they included sound synthesis as one potential application for their data.

More recently this century measurements of a number of instruments were made at TU Berlin as reported by Pollow[10] in 2009, with the suggested application of using these directivity data in room acoustic simulation, such as in the auralisation of a concert hall model. Analysis of many of these measurements was published by Shabtai et al[9] in 2017.

Another data set measured at Aalto University (then Helsinki University of Technology) was reported on by Pätynen and Lokki in 2010[4] with the suggested application of using the instrument directivity data as part of a source model for auralisation, and for further research.

Also in 2010, Hohl and Zotter[15] presented measurements made at IEM Graz on a series of musical instruments with some analysis of the radiation patterns at different pitches.

## 2.2 Sensor arrays and spatial density

The measurement systems outlined in the above section all take a different approach to their sensor array design. The measurement microphones are usually placed equidistant from the notional centre of the array where the instrument being measured is placed. However, as outlined in Deboy and Zotter[16] and in Shabtai et al[9] the acoustic centre of the instrument may not be the geometric centre of the instrument.

Array designs with a low number of microphone positions may use rotation of the instrument (such as the 12-microphone array as reported by Cook and Trueman[3] within the array in order to measure with a higher spatial density of information, so as to have measurement data points that are not physically spaced too far apart. However, this may not work optimally with human actuated instruments due to the inconsistency of human performers. A microphone array with a higher number of microphones such as the 32-microphone array used to generate the TU Berlin data set[8], or the 64-microphone array used for the IEM Graz data measurements[15] will have a higher spatial density of measured data points. This results in having fewer high frequency errors when interpolating between the measured data points, as high frequency radiation may occur in narrow angular lobes which can pass between microphones in the array. While different interpolation methods have been reported such as Hohl and Zotter[15] using spherical harmonic interpolation and Pätynen[17] using linear interpolation, neither method would be able to re-create high frequency radiation that might pass between microphone sensor locations as it would not be captured at all by the measurement sensor array.

### 2.2.1 High spatial density approach

It has been previously reported[6] that a sensor array design based on a 4-frequency expanded icosahedron (which is then truncated to allow for real-world floor-reflection effects as included in the early work by Meyer[14]) with 111 microphone sensors would give an inter-sensor angle of approximately 18 degrees. This array design is shown in Figure 1. This suggests that such an array would present a lesser opportunity for high frequency lobes to pass between microphone positions, increasing the high frequency accuracy of the array over the other designs discussed above.

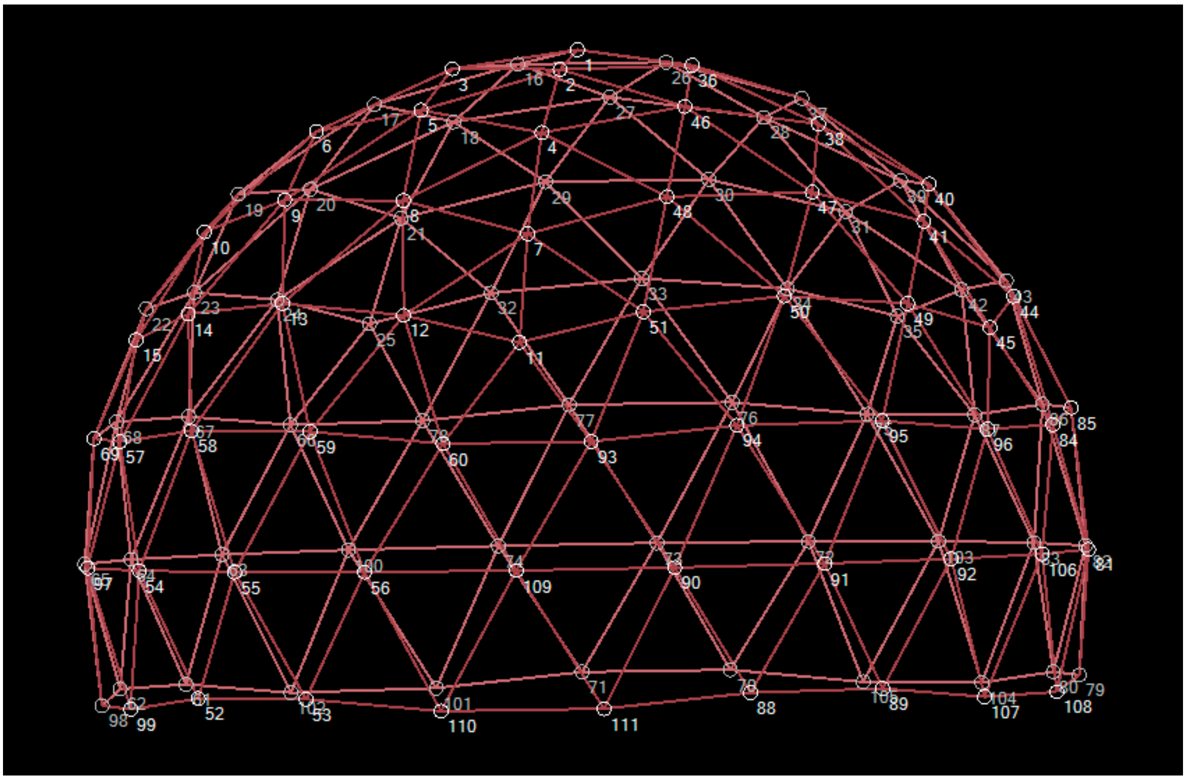

Fig 1. 4-frequency expanded icosahedron with 7/12 truncation showing sensor positions.

# 3 VIRTUAL MICROPHONE SIGNAL PROCESSING

## 3.1 Application of measured directivity data to recorded signals

The process for generating a signal for a virtual microphone from an instrument recording is to take a data set for a musical instrument type, such as a violin or a guitar, or a trumpet, and to use this data to derive a transfer function that could be applied to a recorded signal in a way that applies the spectral characteristic of sound radiated by that instrument in the direction of the virtual microphone position onto the originally recorded signal. This is possible if we know the position of the microphone used to make the recording we wish to process, and if we have a data set for the directivity of the instrument at different frequencies. The result is an output signal that contains the spectral characteristic as if the recording microphone had been placed at the virtual microphone position.

With the acoustic radiation data containing the direction-dependent frequency spectrum, we can compare the spectral content radiated in different directions, and so find the spectral change that we would need to apply to have the spectrum radiated in one direction to sound as if it were actually radiated in the other direction.

In order to generate a signal for a virtual microphone that would give the timbral response as if it were placed in a position chosen by the listener, we can compare the measured data set for that instrument type and find how the spectrum would need to change by using the difference in the measurement data for the original and the virtual microphone positions. Then we would have a transfer function that we can apply to the new recorded signal which would give it the spectral characteristic as if the recording microphone had been placed at the virtual microphone position. The general operation of the signal processing algorithm is shown in Figure 2.

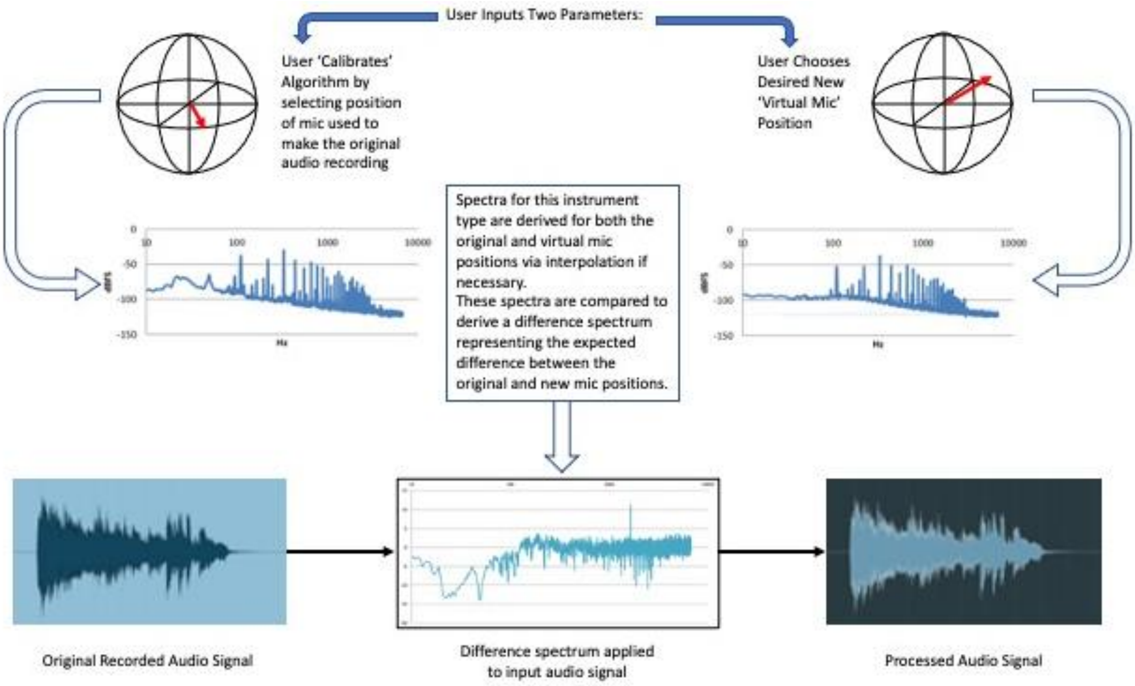


Figure 2. Overview of the virtual microphone signal processing algorithm.

Where the virtual microphone position or the recording microphone position chosen by the user does not correspond to one of the measurement microphone sensor locations in the original data set, we could interpolate in order to find the spectra associated with those required positions.

### 3.1.1 Assumptions

Each individual musical instrument of a particular type will have its own directivity characteristic, and this may be different to other instruments of that type. For example, each violin may be different to other violins. However, directivity behavior for each instrument type can be characterised by taking average measurements for several of that type of instrument, as described by Meyer in 2009[18] and Bodon in 2016[19]. When the algorithm is used to generate a virtual microphone signal for a particular instrument type such as violin or flute, etc, then the underlying data set should be the data set for that type of instrument. Using a data set for a different instrument would not give the expected result, although mis-using the algorithm in this way could be done creatively.

### 3.1.2 Target response

The input signal to be processed will already possess a timbral characteristic, as captured by the recording microphone at the location where it was placed for the recording. The purpose of the signal processing algorithm is to amend the original spectrum by the difference between the spectra of the original and virtual microphone positions which are contained in the musical instrument directivity measurement data set. By applying this difference spectrum (as shown in Figure 2, above) to the input signal, it results in a target spectrum which the output signal will exhibit.

## 3.2 Deriving the transfer function

The difference spectrum that we apply to the input signal is derived by dividing the spectrum of the position of the virtual microphone position by that of the position of the recording microphone as found in the measurement data set. This results in a transfer function (the difference spectrum) that can then be applied to the input signal.

# 4 TESTING AND RESULTS

The signal processing algorithm was implemented using the *pure data* audio processing environment for initial testing of the concept. Some of the issues that arose in this implementation such as time averaging the spectra during analysis may have been due to this implementation and may not be relevant if another environment was used (or indeed with a better implementation within this environment).

## 4.1 Limitations

There were only a small number of test measurements available for use in the prototype. This is likely to have affected the quality of the listening tests by not having a variety of instrument types or test recordings to process, and by having a small number of potential virtual microphone positions.

The prototype implementation described in this paper uses only 1 dynamic level for the test data set. This was considered sufficient for a proof-of-concept, but switching between directivity data sets measured at different dynamic levels could be done by selecting an appropriate quiet-medium-loud dynamic measurement data set as appropriate for the level of the input signal to be processed.

## 4.2 Test data

Although over the course of this work a high spatial density microphone sensor array has been proposed[6], the first implementation of the algorithm used a smaller 7-sensor arc array with an inter-sensor angle of 30 degrees as previously reported[20]. The instrument for which the data set was measured was a Lowden Model O-35 steel string acoustic guitar.

## 4.3 Implementation

### 4.3.1 Directional Response

The algorithm needs input from the user to select the original position of the microphone used to make the recording which is being processed, and to select the position of the virtual microphone. We compare the spectra of radiation in the direction of those microphone positions which gives us a difference spectrum which is to be applied to the input signal spectrum to derive the output signal of the algorithm.

### 4.3.2 Extracting Usable Data

In the prototype implementation described, the analysis of the directional spectra from the data set corresponding to the directions of the original microphone position and the virtual microphone position were done by FFT in blocks of 1024 samples, overlapping each 256 samples. These were divided and the result was time averaged over a number of 250millisecond periods in order to provide a difference spectrum that was stable over time which could then be applied continuously to the input signal to generate the virtual microphone output. The number of averaging periods was judged empirically during the generation of the listening test examples.

### 4.3.3 Applying the transfer function to the test signal

The input signal spectrum is found using an FFT in blocks of 1024 samples, overlapping each 256 samples. The magnitude spectrum is multiplied by the derived difference spectrum (the expected transfer function between the measured original microphone position and the virtual microphone position) and the result is then scaled, after which an inverse FFT to give the output signal.

## 4.4 Testing

In order to get unbiased confirmation that the signal processing algorithm was working as planned, a small group of listeners was recruited from a set of academic and professional colleagues to listen to examples of the processing algorithm and evaluate the result. The people approached for this were considered to be experienced expert listeners with the aural acuity required to make a determination of timbral quality.

### 4.4.1 Listening test

The listening test was planned to be self-administering in that the listeners were given 5 listening examples, where each example allowed them to determine if the processed sound had the timbral characteristics of the original microphone location or the ‘virtual’ microphone location.

The test was conducted twice for each listener: the first time the sample was a single note where the transfer function was generated from a data set based on a measurement of the instrument playing a chord; the second time the sample was a on a chord where the transfer function was generated from a data set based on a single note measurement.

The format of the test was an ABC type test where the listener can hear the 3 examples and must determine if C is more similar to A or B. The audio clip A and B was randomized for the listener and they did not know for each example which was the original microphone position and which was the target microphone position. In all cases the listener did know that C was a signal processed audio clip which was neither A nor B but should have a similar timbral character to one or other of A or B.

The listeners were able to compare C to A and B as many times as they required without a time limit in order to make their determination. They were required to choose C as being more similar to either A or to B; there was no option for neither A nor B and no option for equally similar to both A and B.

### 4.4.2 Listening test results and discussion

For the 5 examples given, the listener determined the processed (i.e. ‘virtual microphone’) signal to either sound more similar to the original microphone position or to sound more similar to the target microphone position in terms of timbre. Each listener undertook this for both the single-note test and the chord test. The results are shown in Figure 3, which include a mean average response for all listeners.

| | Single-Note Test | | Chord Test | |
|---|---|---|---|---|
| Listener | Original | Target | Original | Target |
| 1 | 3 | 2 | 3 | 2 |
| 2 | 5 | 0 | 0 | 5 |
| 3 | 0 | 5 | 5 | 0 |
| 4 | 0 | 5 | 5 | 0 |
| Average | 2 | 3 | 3.25 | 1.75 |

Figure 3. Anonymised listening test results with mean average responses shown.

The results of this listening test could be interpreted in different ways.

In the case that the listener determines that the processed signal is more similar to the original signal then the processing would appear not to be operating correctly or optimally.
In the case that the listener determines that the processed signal is more similar to the target signal, then the processing would appear to be operating as expected.

The results show that for the single note test, 2 of the listeners found that the processed clip sounded timbrally more similar to the target signal in each of the 5 examples. 1 listener found that the processed clip sounded timbrally more similar to the original signal in each of the 5 examples, and 1 listener found that in 3 of the 5 examples the processed clip sounded timbrally more similar to the original signal, and in 2 of the examples the processed clip sounded timbrally more similar to the target signal. Averaging these results shows that the processed clip was found to be more similar to the target 12 times and more similar to the original signal 8 times. While this result is not overwhelmingly positive, it shows some promise and suggests further investigation and refinement may be worthwhile.

However, the results for the chord test are not particularly positive. It is interesting though that for individual listeners the results of the chord test examples are almost the opposite of the single note test: for example the listeners who considered the processed single note to sound more similar to the target signal found that the processed chord sounded more similar to the original clip, and vice versa.

# 5 LIMITATIONS OF THE WORK

Given the restricted nature of the testing done, the prototype described in this paper could be considered as a proof-of-concept rather than a general-purpose implementation. Some specific limitations of this prototype are covered in the following sections.

## 5.1 Target Spectrum Generation

The version of the prototype used for the listening tests used a crude spectral extraction method – an FFT measurement was taken and averaged over a short time within the audio clip from that measurement microphone position within the sensor array. This averaging time was not chosen consistently, but an element of subjective judgement was used to determine the point where the spectrum had settled. Later versions of the prototype have experimented more with the averaging time to generate the spectrum via FFT.

### 5.2 Non-dynamic response

The measurements taken and used for the small data set within this prototype were taken at one dynamic level only. Some data sets such as those measured at TU Berlin include louder and quieter dynamic levels[8] which will affect the measured spectrum of the instrument. These different dynamic levels could be selected depending on the average level of the input signal at any given time.

### 5.3 Listening tests

There was no control test used to calibrate the listener response accuracy. This omission was only realized after the listening tests had been completed, and might have shown if a listeners response was an outlier or perhaps might show that they could not identify the same audio clip in an ABX type test. The number of listening test participants could not be considered statistically significant. Even so, the listener test result for processing of chord sounds was not positive, suggesting that the algorithm as implemented for the listening test had problems in processing this type of signal (whereas the single-note processing was verging towards positive but needing more work).

## 6 CONCLUSION

### 6.1 Summary

The virtual microphone signal can be generated by applying a transfer function to the recorded signal, the target spectrum being derived from a difference between the general instrument response from original recording microphone position and the general response from the virtual microphone position. This difference is then applied to the recorded signal, resulting in the signal obtaining the target spectrum.

### 6.2 Further work

#### 6.2.1 Improvements to implementation and testing

The measurements made for testing the prototype were of a medium dynamic level. Including quieter and louder soundings of the instrument would yield a different spectrum for each dynamic level and could be incorporated into the algorithm by taking a short-time average level of the signal being processed, and selecting over this short time period which target spectrum to use in the processing, i.e. quiet, medium or loud.

The target spectrum extraction method used for the prototype algorithm implementation was not optimal. Modification of the prototype since the implementation used for the listening tests has yielded a seemingly more consistent target spectrum extraction, based on a spectrum capture averaging for audio clip captured at the measurement microphone position.

The number of listening test participants could not be considered statistically significant. However, the aim was not a statistical analysis but to get unbiased expert listeners to offer informed opinions. With further refinements to the algorithm and its implementation, a larger-scale listening test would be appropriate.

#### 6.2.2 Applying the method to other instrument types

Other instrument types could be processed using existing data sets which are described in other research. While this work has used a steel-string acoustic guitar as the test instrument (and as the instrument type from which the test data set was generated in order to generate the instrument type response for the original and virtual microphone positions) it is possible to use other instrument data sets (such as those described by Pollow in 2009[10] and Shabtai et al in 2017[9] to derive appropriate transfer functions for an original and a virtual microphone position for a recording of that instrument type to be processed using this method.

Ideally the processing algorithm could be optimized for different dynamic levels, and a high spatial resolution data set for each instrument type, perhaps averaged over multiple source instruments to be measured as suggested by Meyer[18], would allow the user to select the type of instrument and the appropriate data set measurements would be used in sourcing the original and target spectra for the selected virtual microphone position.

# 7 REFERENCES


1. Meyer, J., Wogram, K., Die richtcharacteristiken des hornes. Das Musikinstrument. Heft 6, 1969, Sieten I-XII (1969).
2. Meyer, J., Wogram, K., Die richtcharacteristiken von Trompete, Posaune und Tuba. Das Musikinstrument 19, 1970, S.171 (1970).
3. Cook, P., Trueman, D., A database of measured musical instrument body radiation impulse responses, and computer applications for exploring and utilizing the measured filter functions. Proceedings of the International Symposium on Musical Acoustics 1998. (1988).
4. Pätynen, J., Lokki, T., Directivities of symphony orchestra instruments. Acta Acustica united with Acustica, vol. 96, no. 1, pp. 138-176, January/February 2010. (2010).
5. Pelzer, S., Pollow, M., Vorlander, M., Auralization of a virtual orchestra using directivities of measured symphonic instruments. Société Francaise d'Acoustique. Acoustics 2012 conference, April 2012, Nantes, France. (2012).
6. Carugo, D., Hayatleh, K., Lidgey, J., Sharp, D., Sensor grid design for high resolution 3D acoustics measurements of musical instruments, Proc of the Institute of Acoustics ACOUSTICS 2019 Conference, Milton Keynes, (2019).
7. Campbell, D.M., Non-linear dynamics of musical reed and brass wind instruments. In Contemporary Physics vol 40(6) pp 415-431. (1999).
8. Behler, G., Pollow, M., Vorlander, M., Measurements of musical instruments with surrounding spherical arrays. Société Francaise d'Acoustique. Acoustics 2012 conference, April 2012, Nantes, France. (2012).
9. Shabtai, N.R., Behler, G., Vorlander, M., Weinzierl, S., Generation and analysis of an acoustic radiation database for forty-one musical instruments. Journal of the Acoustical Society of America, 141(2) pp1246-1256. (2017).
10. Pollow, M., Measuring directivities of musical instruments for auralisation. NAG/DAGA 2009. International Conference on Acoustics, Rotterdam, 23-26 March 2009. (2009).
11. Ismail, S., Gaydecki, P., Lloyd, T., Johannsson, H., Real-Time Emulation of the Acoustic Violin Using Convolution and Arbitrary Equalization. Journal of the Audio Engineering Society, Vol 65, No.5, May 2017. (2017).
12. Pedersini, F., Sarti, A. and Tubaro, S., Object-Based Sound Synthesis for Virtual Environments-Using Musical Acoustics. IEEE Signal Processing Magazine, 17(6). (2000).
13. Pätynen, J., Lokki, T., Concert halls with strong and lateral sound increase the emotional impact of orchestra music. Journal of the Acoustical Society of America 139(3), March 2016. pp1214-1224. (2016).
14. Meyer, J., Akustik und musikalische auffuhrungspraxis. Frankfurt: Verlag Das Musikinstrument. (1972).
15. Hohl, F., Zotter, F., Similarity of musical instrument radiation-patterns in pitch and partial. DAGA 2010 – 36th German Annual Conference on Acoustics, March 15-18, 2010, Berlin, Germany. (2010).
16. Deboy, D., Zotter, F., Acoustic center and orientation analysis of sound-radiation recorded with a surrounding spherical microphone array. Proceedings of the 2nd International Symposium on Ambisonics and Spherical Acoustics. May6-7, 2010. Paris, France. (2010).

17. Pätynen, J., A virtual symphony orchestra for studies on concert hall acoustics. Doctoral Thesis, Aalto University. (2011).
18. Meyer, J., Acoustics and the Performance of Music, Springer, 2009, New York. (Section 4.1.2 P130). (2009).
19. Bodon, K.J., Development, evaluation, and validity of a high resolution directivity measurement system for played musical instruments. Master's thesis, Brigham Young University. (2016).
20. Carugo, D., Hayatleh, K., Lidgey, J., Sharp, D., 'That was Great, Can You Play it Again?': Making 3D acoustic measurements of instruments under performance conditions, In Proceedings of the Institute of Acoustics Reproduced Sound 2015 Conference, Moreton in-the-Marsh. (2015).