Affordable Speaker Arrays

Fernando Lopez-Lezcano*
CCRMA, Stanford University
Stanford, CA, United States
nando@ccrma.stanford.edu

Michael R. Mulshine†
CCRMA, Stanford University
Stanford, CA, United States
mulshine@stanford.edu

ABSTRACT

We present an affordable, lightweight, and highly portable multichannel audio solution for surround sound applications and installations. The system was developed for the "Sound in Space" course, taught by on of the authors at CCRMA in the winter quarter of 2021, when education was fully remote. Students in the course were able to listen to and create surround sound compositions from their homes or dorm rooms. Beyond the course, artists have demonstrated the versatility and creative affordances of this cheap, lightweight, and highly portable setup in sound installations and other custom speaker arrays. Such an affordable and versatile system has the potential to provide more students and artists access to spatialized sound production and multichannel audio in their work, enabling deeper technical education and creative applications ranging from Ambisonics to sound installations. Importantly, the transportability and ease of assembling this system enables multichannel audio work to be developed outside of the physical confines of academic institutions, including in spaces like apartments, garages, the outdoors, and more. This paper steps through the process of creating such a system, detailing the challenges faced and reflecting on the affordances in educational and creative usage.

Author Keywords

Ambisonics, speaker arrays, sound installations, affordability, accessibility, remote education

CCS Concepts

•**Applied computing** → **Sound and music computing;** Media arts; •**Hardware** → Sound-based input / output;

*Designed the affordable speaker array outlined in this paper.

†Creatively repurposed the affordable speaker array in sound installations and other projects. Organized and wrote majority of the paper.

NIME'23, 31 May–2 June, 2023, Mexico City, Mexico.

Figure 1: Left: An affordable at-home 5.1 speaker ring array. Right: An installation built with the same audio hardware.

1. INTRODUCTION

Multichannel audio applications are inherently expensive, requiring large quantities of speakers, amplifiers, audio interfaces, and rigging. In general, the affordability of speakers and amplifiers relates inversely with size and weight, making the highest quality multichannel systems large, heavy, and expensive. More affordable systems are smaller and lighter, making them more portable and flexible. Historically, multichannel audio applications have been developed, tried, and tested in the confines of academic institutions and well-endowed music studios. Access to these spaces is limited for the majority of people.

CCRMA prides itself in making available several high-quality speaker arrays in their facilities[1][11][10], enabling students and artists to easily create multichannel works. However, during the critical months of the COVID pandemic (March 2020 to August 2021), these facilities were closed. Teaching was done remotely, including the bi-annual "Sound in Space" course, taught by Fernando Lopez-Lezcano. The course is dedicated to spatial audio applications and traditionally involves hands-on use of the speaker arrays available at CCRMA. To provide a comparable experience to a remote cohort of students, Lopez-Lezcano investigated a variety of hardware options that were small, lightweight, easily configurable, and, critically, affordable enough to ship to students. This effort resulted in a speaker array system that cost, at the time, only $320-360 per student and was funded by the department. Each student was able to install a 5.1 or 7.1 speaker ring in their home or dormitory, with parts largely sourced from "surround sound" speaker products for movies.

In addition, a suite of open source free software custom plugins was provided to correct for the speakers' frequency response and provide basic tools. A matching workflow that relied heavily on free software tools was prescribed,

[1]`https://ccrma.stanford.edu/`

enabling students to experiment, prototype, and compose from home.

Later, other artists at CCRMA have repurposed the system for use in DIY sound installations in unique environments away from CCRMA, and in custom speaker array development. Despite sacrificing audio fidelity and power handling, we show that many useful and creative multichannel audio applications are worth pursuing without the "best" equipment, and in fact, many applications are enabled by this tradeoff, including remote education.

The speaker array system explained in this paper here is by no means state-of-the-art in terms of fidelity, but with some fine-tuning, performs well for beginner spatial audio education and creative prototyping. While we are writing about particular brands or models, we are not endorsing any one of them as a best option. What is described is our approach to tuning an affordable and viable speaker array, undertaken in the summer of 2020 while preparing to teach the "Sound in Space" course. Available hardware and pricing changes over time, so it would be advisable to re-evaluate the choices described in this paper, as different alternatives may be found now or in the future.

2. DESIGN AND IMPLEMENTATION

A minimal surround speaker setup involves a ring of 6 to 8 speakers spaced regularly around the listening position. In addition to the speakers, an audio interface, amplifiers, cabling, and speaker stands were needed. Finding affordable hardware within the limited funding available was crucial, as we wanted all students enrolled in the course to have such a setup at home.

2.1 The Search for Affordable Hardware

Initially we considered using multiple very cheap stereo speakers meant for computer desktop applications but this required additional cables, power strips, and adapters, increasing the price. We started exploring small commercial surround speaker systems and found one by Logitech which could be purchased for only $130 at the time. The product included five small satellite speakers, a combined subwoofer and power amplifier, plus all the required speaker cabling.

High quality multi-channel audio interfaces have become quite affordable, but were too expensive for the course budget. A search for options revealed the Startech 7.1 USB audio interface (or its clones). These cost only $35. They provide four stereo 1/8-inch unbalanced outputs, 16-bit DAC's, and a 48KHz maximum sampling rate. While balanced outputs, 24-bit DACs and 96kHz sampling rates are the norm, these specifications were good enough for our low-budget needs.

To mount and position the speakers we looked for cheap speaker stands that could fit in small spaces at home or in dormitories, but even the cheapest cost more than $100 a pair. DIY options were considered, including custom PVC pipe statnds, but the costs and labor added up. We finally found small stands meant for photography lighting that cost between $12 and $20 each. As no affordable solution existed for attaching the speakers to the stands, we designed a simple adapter that could be easily 3D-printed (See Figure 2).

2.2 Initial Tests

At this point, the larger hardware problems were solved, but we had not yet worked to improve the audio quality of the speakers. Preliminary frequency response measurements were performed using JAPA[4], a realtime percep-

Figure 2: Left to right: combined sub and amplifier unit and 5 speakers, 7.1 channel audio interface, photography light stand, threaded mounting adapters.

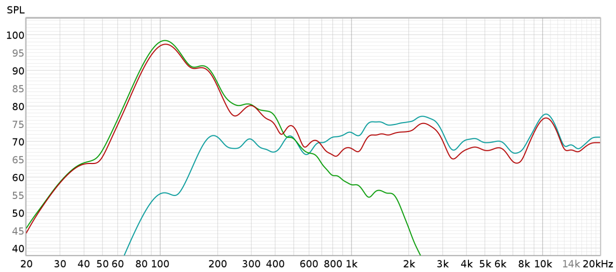

Figure 3: Frequency response of the 5.1 system without any calibration, measured with REW. Green trace is the subwoofer, blue trace the left speaker and red trace both combined

tual spectrum analyzer, and a calibrated microphone (the Dayton EMM-6 Measurement Microphone). Later measurements used the free Room Equalization Wizard software (REW[12]) (see Figure 3). These tests showed that this particular system was far from ideal, as expected. The subwoofer's response featured a single substantial peak at around 90Hz, followed by a gentle rolloff at higher frequencies. The satellites exhibited uneven frequency response, with a wide dip between 3 and 8.5KHz.

The question became: could this system be tuned to provide adequate performance for basic listening of surround pieces for the course, and for creation of new content by the students?

2.3 Tuning the Speakers

CCRMA's concert speaker array ("the GRAIL,") and associated control hardware, has been in use for live diffusion in concerts since approximately 2011. Around 2014, we set out to improve the sound quality, with the goal of bringing studio quality sound to our multichannel concerts[9]. Inspired by a paper[13] presented at the 2008 Linux Audio Conference, we began using the DRC software (Digital Room Correction[14]) to equalize each speaker individually, with very good results.

A similar approach was taken for the speaker array described in this paper. We aimed to provide students an easy-to-use plugin, compatible with any operating system or audio platform they might choose. Our software platform of choice for this task was the FAUST[2] programming language, which is designed for digital signal processing and can be compiled into nearly every existing plugin format or platform.

The software we were using for the GRAIL (DRC) creates convolution kernels, but convolution is not yet supported by FAUST, so we ended up using REW (Room Equalization Wizard) to design parametric equalization filters to approximate a flat frequency response. We used REW to perform measurements of the subwoofer and satellite speakers deployed in a typical small room, and time-aligned and vector-averaged the 5 main speakers to arrive at an average response, and a set of parametric filters was designed using REW's design tool (four or five sections proved to be adequate). The same process was repeated for the subwoofer. See the results for the combined vector average response of the small speakers in Figure 4.

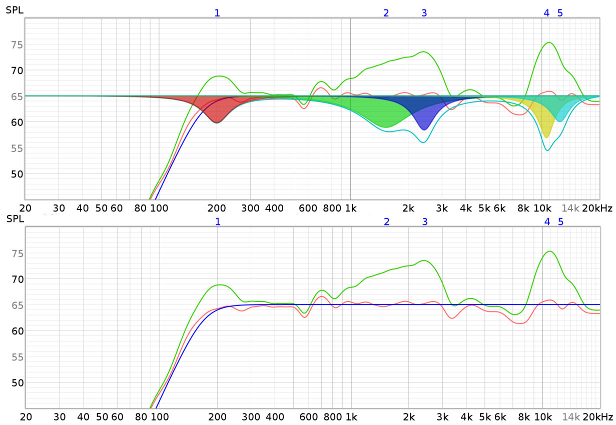

Figure 4: Frequency response of all speakers. The blue trace is the desired response, green is the measured response and red is the corrected response. The upper part of the figure shows the individual and combined responses of the parametric sections.

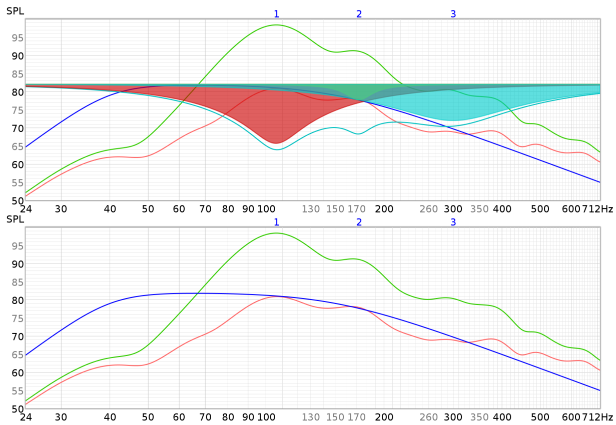

Figure 5: Frequency response of the subwoofer. The blue trace is the desired response, green is the measured response and red is the corrected response. The upper part of the figure shows the individual and combined responses of the parametric sections.

We also designed a phase-aligned crossover, and equalized levels to achieve a largely flat frequency response. All of this was incorporated into a small FAUST program which was then compiled into plugins that could run on all operating systems and audio platforms. Figure 6 shows the frequency response of the whole system, measured with REW's real-time analyzer. The subjective results were very good compared to the original performance of the array. In qualitative assessments, the system processed through the FAUST plugin sounded quite neutral and rendered materials adequately.

The downsides of such an inexpensive system are not limited to very uneven frequency response and low and high frequency coverage limits. Power handling is also limited, especially when coupled with the filters needed to improve the frequency response. As the suggested listening distance for this system was small, this was not deemed a problem. Distortion can also be a problem, especially at higher loudness or when playing content with a wide dynamic range. Despite these tradeoffs the array could deliver a satisfying level of performance for education and newcomers to ambisonics.

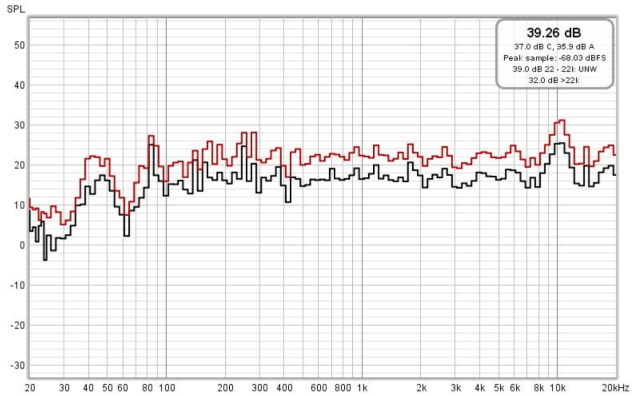

Figure 6: Frequency response of the whole system with crossover and equalization filters measured with the REW realtime analyzer.

2.4 Software

The software developed for the course, all free and open source, included the aforementioned "speaker processor" plugin that hosts the parametric equalizers, the subwoofer crossover, and a level control to balance the level of each speaker.

Additional FAUST plugins were provided for Ambisonics decoders and panners that integrated with Ambisonics reverberation plugins. The decoders were calculated using the Ambisonics Decoder Toolkit (ADT[7][6]), and the panning formulas were created from symbolic spherical harmonic equations generated by a Python software program[5].

We also provided a document that described Ardour[1] (a free software digital audio workstation) and Reaper[2] (not free but reasonably priced) workflows. Additional plugins for Ambisonics, including the AmbiX[8] suite and the Ambisonics Toolkit (ATK) for Reaper[3] were prescribed in tandem with our custom tools. All of this enabled students to download example spatial audio compositions, listen to them at home, and to begin to create content.

The Faust plugins, 3D models for parts, and the DAW workflows are hosted in the still-evolving "Surrounded at Home" GIT repository[3].

2.5 Pricing and Shopping

The complete system as deployed to each student (including the Z606 surround system, the audio interface, additional cables, stands, 3d printed adapters, and software) totaled $320 to $360 per seat at the time the course was taught. A 7.1 system would have costed $450 to $480 per seat.

It cannot be stressed enough that shopping online for alternatives takes a long time, requires patience and is quite frustrating. Scrolling through tens of pages to find the best

[2] `https://www.reaper.fm/`

[3] `https://cm-gitlab.stanford.edu/audio/SurroundedAtHome`

option for one item is not unusual, as online shopping search engines are not designed to give you the best option right away. The same search would find different results on different days, and/or prices would fluctuate wildly, so the authors encourage those with similar aims to do their research as markets evolve.

2.6 Expanding the System

Many possibilities emerge when using more than one Z606 and audio interfaces. By using two Z606 systems we can assemble a 7.1 system (with 3 leftover speakers) that fully utilizes all 8 outputs of a single audio interface, for improved spatial resolution. Operating systems that can aggregate multiple audio interfaces open up the possibility of using more than one audio interface. For example, as some of the stands that were purchased were very tall, we tested mounting speakers "overhead" in a 5.1.4 (four upper speakers) configuration that could render full 3D surround sound. Combining two Z606 and two audio interfaces enables the creation of a 10.2 system for quite high resolution horizontal rendering, or a 7.1.4 system for even better 3D reproduction. Simple tests showed these were viable alternatives for home listening to pieces created for full 3D surround.

3. APPLICATIONS AND EVALUATION

The speaker array solution proposed in this paper has some interesting and unconventional qualities. On the one hand, it's less *about* quality. We intentionally don't employ the best speakers, primarily to reduce costs, but also to reduce weight and increase portability. Smaller and lighter speakers also require less wattage, reducing the cost of amplification. While some may see all of this downgrading as a negative, the authors have found through practice that this solution provided a tremendous amount of convenience in education, as described in the following section (Section 3.1). It also provided a large amount of physical reconfigurability and creative flexibility attractive to sound artists, inspiring unique multichannel audio projects in unlikely spaces and moments, as detailed in Section 3.2.

3.1 Course Feedback

Students in the Sound and Space course were able to experience surround sound at home (see Figure 7), without the usual physical access to our studios. In an anonymous survey, a student noted that the audio quality of the system was "not great" but that "it really gave [a] solid impression of the space, much better than the stereo." Despite the downside of not being able to use CCRMA's speaker systems, there were unique educational benefits to this remote DIY method. One student noted feeling inspired by the many considerations that go in to constructing a viable acoustic system in a space - measuring the distance between speakers, considering flooring (carpet or hardwood), considering the shape of the room and whether to acoustically treat walls - and added, "I actually think I will carry over more skills into the future and feel more confident being able to do something like this again on my own." Working at the institution, they would not have learned about these practical and acoustic considerations in a hands-on manner.

3.2 Creative Uses / Misuses

Author Mike Mulshine has used the core of this affordable multichannel audio setup in a variety of creative works. Notably, Mulshine has leveraged the compactness and portability of the setup to create multichannel audio installations

Figure 7: Two student setups from the remote Sound in Space course.

quickly (on-the-fly) in vernacular spaces, such as apartments and garages. On one occasion, Mulshine and colleague Dirk Roosenburg assembled an 11.2 speaker wall (using the Z606 sub and speakers) suspended by string from drop ceiling in a small apartment (see Figure 8). They used two Z606 sub-amp units to amplify 10 channels and quickly hacked together an additional amplifier (for the 11th channel) using an Adafruit MAX98306 stereo amplifier[4]. The speakers were light enough to hang by string from standard drop ceiling. Mulshine embedded a distance sensor in the center of the array and sonified movement in the space in front of the speakers. Additional Z606 speakers, Vantec interfaces, and contact mics triggering sounds were hidden around the space just outside the apartment, completing a site-specific installation for a party. Visitors interacted with the spatialized sound wall and other hidden sound sources while otherwise enjoying the social moment.[5].

Mulshine took this speaker wall concept further by creating another installation called *near and far*. It featured a 14.3 suspended speaker array hanging from the ceiling of a garage, complete with three Z606 amplifiers, three aggregated Vantec interfaces, and fourteen satellite speakers. The speakers were mounted to a wooden frame with 3D printed threaded inserts affixed. Similar to the previous project, a distance sensor was attached to the center of the speaker array. Three pull-string lamps with photoresistors attached to their bulbs allowed audience members to toggle between 8 different compositions or soundscapes which were modified in real-time by the users proximity to the speaker wall. By giving users the ability to interact with and modify a variety of soundscapes, the installation provided agency to audience members to engage with the musical experience on their own terms. The core of the project was installed in approximately 5 hours and presented at the CCRMA Open House on October 21st, 2022 in a garage space behind the CCRMA facilities (See Figure 9)[6].

Mulshine has also begun constructing custom speakers via

[4]`https://www.adafruit.com/product/987`
[5]Documentation: `https://youtu.be/uF5RVmPYJ_k`
[6]Documentation: `https://youtu.be/02L4c2EdC0k`

Figure 8: ***near and far*** **- an interactive installation assembled with the affordable speaker array outlined in the paper.**

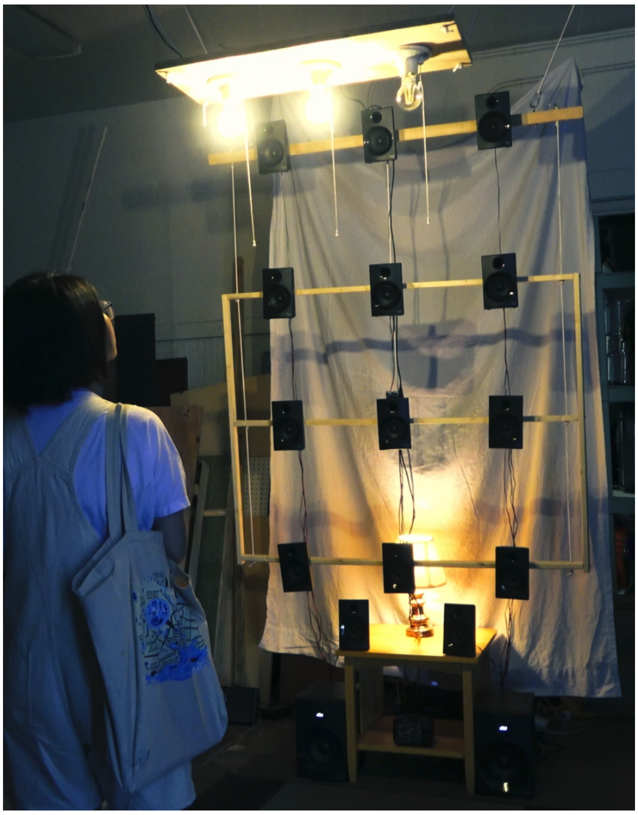

Figure 9: An audience member engaging with the 14.2 interactive sound and light installation, ***near and far***

.

combinations of 3D printing and DIY construction. These include hanging speakers made of PVC for an overhead suspended sound installation, and speakers mounted in pegboard with possible wavefield synthesis possibilities (see Figure 10).

The compact and lightweight nature of the audio solution presented in this paper gave Mulshine the freedom and ease with which to creatively develop these quirky applications. The authors believe that this ease of reconfiguration and portability makes the affordable speaker array solution proposed in this paper (and other similar systems) an effective pathway for artists and creatives to experiment with multichannel audio in their work. With access to such tools, artists can make multichannel work more easily, more often, and outside of the confines of well-endowed institutions, importantly broadening access and possibilities for engagement with multichannel sound projects.

4. FUTURE WORK

The authors will continue to improve the speaker array and explore other affordable multichannel audio production methods. A more expensive (about $400) model, the Z906, has much better speakers, better crossovers, and substantially better sound quality and power handling. After equalizing the speakers using REW a Z906 system yields much better spatial rendering quality. Lopez-Lezcano used a Z906 to add an upper layer to his 8 channel home studio, allowing the creation of works with full 3D surround sound.

High cost per channel of amplification remains the primary limiting factor for the creation of multichannel audio applications (speaker arrays, sound installations), especially outside of institutions. The Logitech system or other equivalent small surround system provide just 5 channels (plus a sub) of amplified signal at 11W to 60W per speaker. Scaling up the quantity of channels would require many more of these amplifier-sub units (increasing costs quickly) or an alternative amplification solution. Efforts to create a multichannel power-amp that provides smaller levels of wattage per channel (anywhere from 3 to 15W) is underway, providing 8 or 16 amplifiers (e.g. the LM386) on a single PCB board design.

5. ACKNOWLEDGMENTS

We would like to specially thank Chris Chafe (CCRMA's Director) and Stanford University for allocating funding that allowed the original project to happen. Many thanks also to Davor Vincze (a CCRMA student) who helped test the first systems during the Summer of 2020 (both hardware and software) providing valuable feedback that helped improve the user experience.

6. ETHICAL STANDARDS

This project was supported by funds provided by CCRMA at Stanford University. Students involved in the Sound in Space course were not treated as or evaluated as subjects, but rather learned about ambisonics and multichannel audio by employing the system devised in a curricular context.

7. REFERENCES

[1] Ardour, a digital audio workstation. `https://ardour.org/`.

[2] Faust, a functional programming language for sound

Figure 10: Left/center: A PVC speaker enclosure with dacron insulation, rubber gasket, and laser cut backing, meant to hang from above by string. Right: 3D-printed speaker enclosure designed to be mounted through pegboard holes, with peg and clip mechanism.

synthesis and audio processing. https://faust.grame.fr/.

[3] ATK, the ambisonics toolkit for Reaper. https://www.ambisonictoolkit.net/documentation/reaper/.

[4] F. Adriansen. JAPA), a perceptual audio spectrum analyser. http://kokkinizita.linuxaudio.org/linuxaudio/.

[5] A. Heller. Symbolic spherical harmonics. https://bitbucket.org/ajheller/symbolic_spherical_harmonics.git.

[6] A. Heller and E. Benjamin. The ambisonics decoder toolbox: Extensions for partial coverage loudspeaker arrays. *Proceedings of the Linux Audio Conference 2014*, 2014.

[7] A. Heller, E. Benjamin, and R. Lee. A toolkit for the design of ambisonic decoders. *Proceedings of the Linux Audio Conference 2012*, 2012.

[8] M. Kronlachner. Ambix: cross-platform ambisonic processors. http://www.matthiaskronlachner.com/?p=2015.

[9] F. Lopez-Lezcano. Searching for the GRAIL. *Computer Music Journal*, 40(4):91–103, 2016.

[10] F. Lopez-Lezcano and C. Jette. Bringing the GRAIL to the CCRMA Stage. In *Proceedings of the Linux Audio Conference 2019*, 2019.

[11] F. Lopez-Lezcano and J. Sadural. Openmixer: a routing mixer for multichannel studios. In *Proceedings of the Linux Audio Conference 2010*, 2010.

[12] J. Mulcahy. Room equalization wizard (REW), free software for acoustic measurements. https://www.roomeqwizard.com/.

[13] J. Nettingsmeier. AMBI@Home the search for extrafrontal intelligence. In *Proceedings of the Linux Audio Conference 2008*, 2008.

[14] D. Sbragion. Drc: Digital room correction. http://drc-fir.sourceforge.net/, 2002.