Acta Acustica united with Acustica
Vol. 94 (2008) 856 – 865

DOI 10.3813/AAA.918104

Anechoic Recording System for Symphony Orchestra

Jukka Pätynen
Department of Media Technology, Helsinki University of Technology, P.O.Box 5400, 02015 TKK, Finland. Jukka.Patynen@tml.hut.fi

Ville Pulkki
Department of Signal Processing and Acoustics, Helsinki University of Technology, P.O.Box 3000, 02015 TKK, Finland

Tapio Lokki
Department of Media Technology, Helsinki University of Technology, P.O.Box 5400, 02015 TKK, Finland

Summary
A method for recording symphonic music with acoustical instruments in an anechoic chamber is presented. Excerpts of approximately 3 minutes were recorded from orchestral works representing different musical styles. The parts were recorded separately one at a time in order to obtain perfect separation between instruments. The challenge was to synchronize different takes and parts so that they could later be combined to an ensemble. The common timing was established by using a video of a conductor conducting a pianist playing the score. The musicians then played in an anechoic chamber by following the conductor video and by listening to the piano with headphones. The recordings of each instrument were done with 22 microphones positioned evenly around the player. The recordings, which are made freely available for academic use, can be used in research on acoustical properties of instruments, and for studies on concert hall acoustics. This article covers the design, installation, and technical specifications of the recording system. In addition, the post-processing, subjective comments of musicians as well as potential applications are discussed.

PACS no. 43.75.-z, 43.55.Ka, 43.58.Ta

1. Introduction

In auralization process an anechoic stimulus is convolved with a modeled or measured impulse response of a space so that the listener can have a binaural listening experience [1]. In general terms, auralization can be seen as a communication process, which is often described as follows— a content, emitted by a source is propagated through the medium to the receiver. Thus, in auralization process an anechoic stimulus signal is the content, a sound source is the source, room acoustics is the medium, and a listener is the receiver.

Of these three parts of auralization process, room acoustics modeling has been under active research for over 40 years [2] and there exist plenty of different modeling methods [3, 4, 5]. Also listener modeling, i.e., studies on human binaural hearing and head-related transfer functions [6, 7], has gained a lot of research interest. However, studies on spatial properties of real sound sources and stimulus signals are seldom found.

Received 29 February 2008,
accepted 1 October 2008.

As a symphony orchestra is the most common sound source in concert halls, it should also be used when auralizing acoustical models of halls. In auralization, each instrument should be simulated to different position on the stage, and also the signal emitted by the instrument to different directions should be available. Thus, the need is to record the orchestra in a way that the sound from each instrument is captured to individual microphone channels, preferably with many microphones around the source in 3-D. In addition, each microphone should capture only one instrument, which is generally not possible, if the instruments are played simultaneously in the same space.

Only few recordings of anechoic orchestral music are published [8], of which the most accessible is by Denon [9, 10]. Unfortunately this recording is quite noisy and the whole orchestra has been recorded at the same time with close-up microphones inside a sound absorbing shell constructed on a concert hall stage. With this technique different instruments are not perfectly separated in microphone channels due to the crosstalk in the setup. In addition, floor reflections occur in larger magnitude than in specific anechoic laboratory conditions. Some other recordings have also been made for research purposes. A choral ensemble has been recorded in an anechoic chamber [11]. Otondo

Table I. List of the musical terms used in this paper.

crescendo	gradual change of the volume to louder
divisi	passage where an instrument section divides to play multiple parts instead of a single one
fermata	pause of longer duration than the written note value
fortissimo	as loudly as possible
mezzo-piano	moderately quiet
répétiteur	a pianist replacing an orchestra, e.g., in opera rehearsals
tempo	overall speed of music
tremolo	style of playing notes in rapid repetitive manner
tutti	passage where the whole orchestra is playing at the same time
unisono	passage where multiple instruments are playing exactly the same notes

and Rindel [12] have conducted investigations on solo instrument directivities with 13 microphones. Recently, Vigeant *et al.* [13] have applied recordings of an orchestra in multi-channel auralization, but besides the number of used microphones, the recording process of anechoic stimulus material has not been reported.

Information about directional characteristics of musical instruments has been presented by Meyer [14, 15], and some numerical data on these investigations is available in the Internet [16]. In addition, Fletcher and Rossing [17] discuss about directivity of instruments. However, a complete data set of directivities of the instruments of a symphony orchestra has not been presented with detailed documentation of the recording process.

The objective in this project was to record orchestral music in an anechoic chamber one instrument at a time in order to obtain symphony orchestra music with perfect channel separation. Another aim was to perform the recordings with multiple microphones for later instrument directivity analysis. In individual recording of instruments, a major challenge was to provide information on synchronization for musicians so that they could play as an ensemble with common timing and tuning. In this paper a method and implementation for such a recording system is described. The recorded audio tracks are made available free for academic use.

The paper is organized as follows. The design and installation of the recording environment are explained first. In Section 4, recorded music is described with the preparation of the conductor video. In Section 5 the actual recordings are described with comments of the musicians. The post-processing of the recorded music is explained in Section 6. Finally, applications for recorded sound tracks are introduced before the conclusions. The musical terms used in this paper are collected in Table I.

2. Recording technique

The instruments were recorded one by one in an anechoic chamber. After discussions with a few professional conductors the following way was applied to enable the synchronization. The musicians played their parts with the help of a conductor video by watching a conductor in a monitor and simultaneously by listening to a pianist playing the whole score, see Figure 1. This way the musicians were able to adapt their playing style and tempo, and the synchronization between different players was possible. Thus, providing the video and audio tracks to the musicians had to be taken into account in the system design.

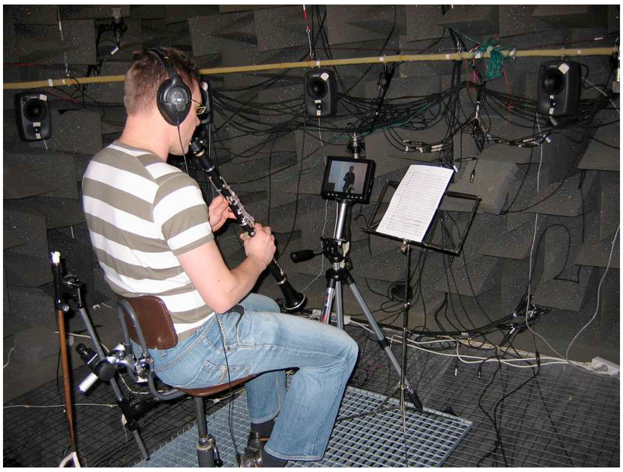


Figure 1. The recording configuration in an anechoic chamber. A musician followed the conductor through the monitor and listened to the piano version of the whole score while playing his/her own part.

3. Design and recording equipment

3.1. Anechoic chamber

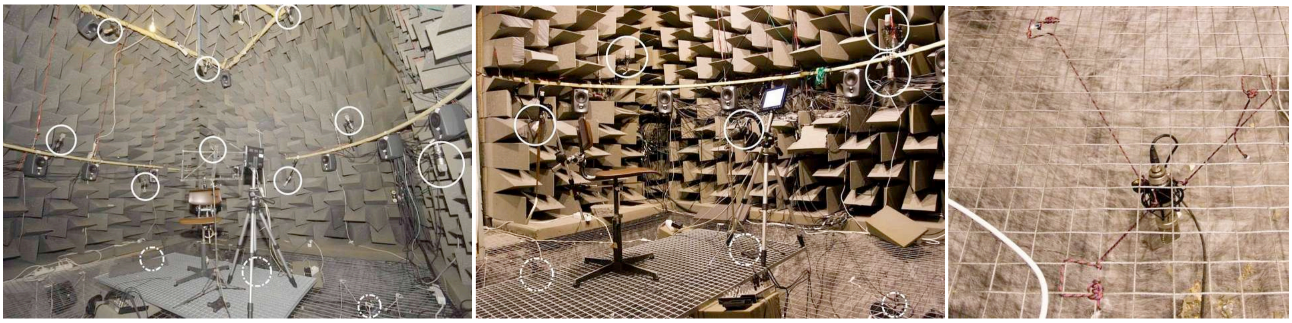
The anechoic chamber used for the recordings is cubical, and the free measure between wedge tips in each dimension is 4.2 m. With the absorption wedge length of 80 cm, the room is assumed to be anechoic at frequencies above 100 Hz.

Anechoic conditions are required for two reasons. First, standing waves affect to the frequency response in different microphones positions. Due to the room modes and limits of anechoic conditions, the measurements below 125 Hz are considered approximate. This issue is more pronounced in investigations on instrument directivity.

Second, the recordings should be free of room response i.e., reverberation. However, the recorded instruments producing fundamental frequencies significantly below 100 Hz are timpani, tuba, contrabass, and cello. These instruments have a noticeable decay time, thus reducing the importance of anechoic conditions. Therefore the absence of completely anechoic environment at low frequencies is not considered as a major problem is this context.

3.2. Microphones

For recording sound in multiple directions, 22 Røde NT1-A-type large-diaphragm microphones were installed to the anechoic chamber. According to the manufacturer data,



(a) (b) (c)

Figure 2. Examples of microphone positions in the anechoic recording system.

Table II. Elevation and azimuth angles, distances and used numbering of microphones. Distances from the center of the room is denoted with r. Note that microphone no. 14 was aligned off the intended position due to the doorway.

Microphone	Elevation [°]	Azimuth [°]	r [m]
1	52.6	0	2.43
2	52.6	72	2.24
3	52.6	144	2.46
4	52.6	216	2.49
5	52.6	288	2.49
6	10.8	0	2.30
7	10.8	72	1.94
8	10.8	144	1.92
9	10.8	216	2.14
10	10.8	288	2.25
11	-10.8	36	2.16
12	-10.8	108	2.03
13	-10.8	180	1.87
14	-10.8	249	1.81
15	-10.8	324	2.06
16	-52.6	36	2.05
17	-52.6	108	2.04
18	-52.6	180	2.00
19	-52.6	252	1.92
20	-52.6	324	2.08
21	0	0	2.21
22	90	0	2.06

NT1-A microphones feature a low self-noise ($L_{\text{noise,A}}$ = 5 dB).

Twenty microphones were geometrically positioned to form a shape of a dodecahedron. The shape was selected due to the equal distances between adjacent vertices. In addition, the microphones positioned at the vertices form a rough representation of a spherical surface. The dodecahedron was oriented to form four horizontal microphone levels, each consisting of five microphones in a regular pentagon. Besides the microphones in the dodecahedron shape, two additional microphones of the same kind were positioned to the front and above directions from the center point. The numbering, angles and distances of microphones are presented in Table II.

The center of the room was defined with an optical measurement device. From this point, the microphone positions were measured by using a laser pointer attached to a theodolite. This procedure allowed the directions to be adjusted to ±1 degree accuracy. As an exception, the microphone no. 14 at the doorway had to be positioned slightly off the correct position to ensure a more secure location. The actual microphone mounting was accomplished with telescope booms that were attached to the top frame (top microphones circled in Figure 2a), the middle rim (Figure 2b) or suspended from the floor net (Figure 2c). Note that the loudspeakers seen in Figures 1–3 and 5 were not used in the recording system, but they were in the room for other purposes.

The distances from the center of the room to each microphone were measured with a laser meter. They were between 1.81 and 2.49 m, while the average distance was 2.13 m. Top level microphones were the furthest from the center, the average distance was 2.42 m. The microphones at the two middle elevation levels were positioned as far from the center of the room as possible, still avoiding the proximity of the tips of the absorption wedges to the microphones.

More accurate microphone positioning was hampered by the fact that we refrained to position the microphones close to the loudspeakers in order to prevent coloration in recordings. The loudspeakers were at least at the same radius from the center of the room as the microphones. Hence the effect caused by the loudspeakers was predicted mostly as diffraction instead of reflections from loudspeaker cabinets.

Besides the flexible steel net shown in Figure 2c, the room featured an existing 1×1 m rigid steel grid as an acoustically transparent floor. This material is similar to the grid used in metal stairs. In order to accommodate larger instruments, such as timpani, an additional 2 m^2 steel grid visible in Figures 1, 2 and 3 was installed on the top of rubber dampers to prevent any rattling between the steel parts. Although the larger grid was predicted to cause some scattering, it was needed to provide sufficient support for the musicians.

The differences between responses of individual microphones as well as the effect of the objects in the anechoic chamber were compensated with filters designed with equalizing measurements. This process is discussed in detail in later sections.

The defined center position was selected as the position of the head of the musician during recordings. Another possibility would have been the acoustical center position of each instrument. However, the acoustical center for each instrument would be difficult or even impossible to define. Exceptions from the head position were made in three cases, where the instrument is commonly played other than sitting: contrabass, percussions and singing. A soprano singer performed in standing position like in a real performance situation. Hence the head, now considered also as the main source of radiation, was 40 cm higher than with other recorded instruments. In spite of this difference, both cases represent actual positions on a concert hall stage. For the largest recorded instrument, timpani, the measurement results are predicted to be only approximate due to the close distance to the microphones, thus the re cording is not performed in the far field. The recording situation for timpani is depicted in Figure 3.

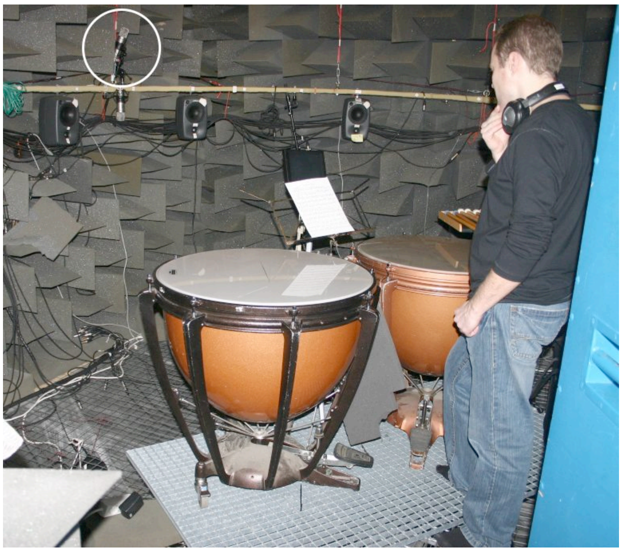


Figure 3. Positioning of the timpani in anechoic chamber. Microphone no. 6 in forward direction is circled.

Figure 4. From bottom to top: Presonus 8-channel microphone pre-amplifiers; Motu 2408 mk3 audio interface; Soundcraft mixer for monitoring.

3.3. Recording hardware and software

For the recording purposes an Apple Mac Pro computer was used running a single patch programmed within Max/MSP environment. The patch was designed to play the video from the beginning until the end, and record the instrument on 22 microphone channels and piano track on one channel during that time period. The piano track was recorded to ensure perfect synchronization in post-processing.

The computer was equipped with Motu 2408 mk3 audio interface featuring both analog and ADAT optical I/O connections. The microphones in the anechoic chamber were routed through three 8-channel Presonus Digimax FS pre-amplifiers to the Motu 2408. This part of the setup is shown in Figure 4. Two remaining channels were used for a talk-back microphone and a piano track loopback to the recording patch.

Audio monitoring outside the anechoic chamber was organized by routing a signal selectable from any of the microphones to a monitoring mixing console along the piano track of the conductor video. Thus, it was possible to monitor the performance against the piano track while following the score.

While playing in the anechoic chamber, the musician wore open headphones (Sennheiser HD-590) for listening to the piano track of the conductor video as well as for communication. Instead of closed headphones, open ones were used in order to help the players hear their playing without additional monitoring. A potential problem with open headphones is that they radiate sound also outwards. However, despite this, the headphone signal did not leak to the microphones audibly. In case a player would have needed to hear more of his/her own playing, one selectable microphone channel was routed optionally to the headphones. After all, nearly all players preferred to wear headphones over only one ear. Despite of the unnatural acoustic conditions, self-monitoring was used only once. Similarly to studio recording convention, providing monitoring signal to the musician with added reverberation was planned prior to the recordings. However, this was seen as potential source for inaccuracies in synchronization and it was not implemented in the recording system.

Since the longest recording had the duration of 3 min 47 sec, recording 23 tracks at 16 bit and 48 kHz sample rate took more than 500 MB per take. Therefore, recording several takes per part consumed unusually large amounts of hard disk capacity. In total circa 70 Gbytes of data were gathered.

3.4. System equalization

The accurate analysis of the directivity of the instruments requires that the responses of the microphone channels have to be equal. Therefore the sensitivity and magnitude response of each microphone had to be measured in their final positions due to the acoustical differences between individual microphones, amplifier channels and positions in the room. Specific filters could then be designed to equalize the responses of the recording channels for post-processing phase.

Suitable recording levels were sought by hitting a snare drum and cymbals in the recording room. Sound levels from such percussion instruments were regarded to represent the maximum recorded levels. The gain control buttons on the microphone amplifiers were fixed to the same position for the duration of the whole recording project. As a downside, the dynamic range could not be fully utilized, since many instruments were considerably quieter than the loudest ones.

The actual calibration and equalization process of the recording system is described in the following.

First, the Genelec 1032A [18] loudspeaker used for the calibration process was measured in an empty, large anechoic chamber with one Brüel & Kjær 4191 measurement microphone and with a few Røde microphones. Being considerably larger than the instrument recording chamber, this anechoic chamber has been measured to be anechoic at frequencies above 80 Hz. The B&K response is considered ideal. The measured microphones were aligned with a laser meter attached on top of the loudspeaker cabinet (see Figure 5). The laser beam was then pointed to the band next to the microphone grill. In the large anechoic chamber the measuring distance from the loudspeaker was 2.13 m, which is the same as the average distance of the microphones from the center of the recording room (see section 3.2). A sine sweep [19] was used to measure impulse and frequency responses.

The measurement results obtained in large anechoic chamber are depicted in Figure 6, where the response of the B&K is with dashed line and the response of one Røde with solid line. The peak visible at 60 Hz is a previously known feature in the large anechoic chamber. The boost at 1-3 kHz in B&K response results from the crossover frequency of 1.8 kHz of the reference loudspeaker, since the microphone was not on axis between the elements of the loudspeaker, as shown in Figure 5. The most apparent feature in the Røde microphones is the pronounced response at high frequencies.

Second, the responses of the 22 Røde microphones in their final positions in the recording room were measured. For this purpose we used the same Genelec 1032A loudspeaker, which was mounted on a tiltable and rotatable stand at the center of the room. Again, the attached laser meter was utilized to align the speaker towards the microphone during the response measurements and to measure the distance. The frequency responses from these measurements are shown in Figure 7 (B&K response in the large anechoic chamber plotted with dashed line). The response of the loudspeaker is present in all responses, showing similar overall behavior as in Figure 6.

The objective in the equalization was to compensate the differences in sensitivity and frequency response of the

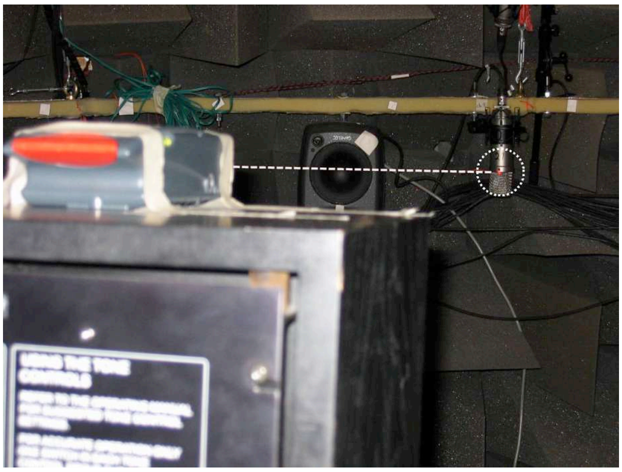


Figure 5. Illustration of the microphone measurement. The loudspeaker is targeted towards a microphone for a response measurement. The bright red dot at the center of the drawn circle is from the laser distance meter attached on top of the measurement loudspeaker.

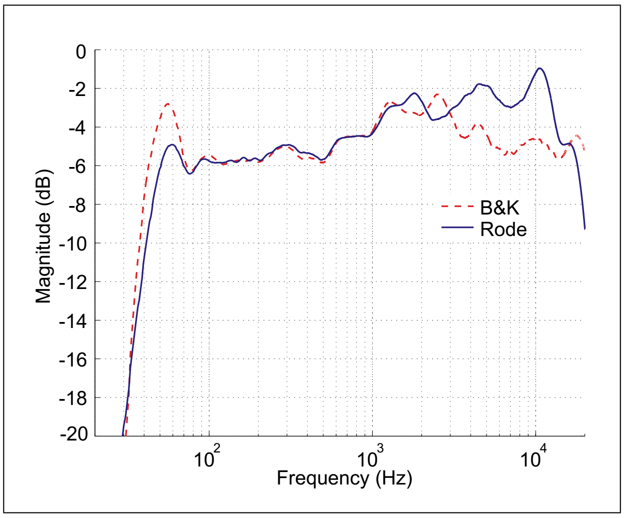


Figure 6. Comparison of B&K (reference) and Røde (recording) microphone responses with measurement loudspeaker response. Responses are smoothed to 1/3-octave resolution.

Røde microphones compared to the B&K response. Based on the measurements, it was possible to compute equalization filters for each of the 22 Røde microphones. For the filter design, specific target responses were required. This was accomplished by deconvolving the Røde measurements in the recording room with the B&K measurement. These target responses were then utilized in designing the actual equalization filters for the recording microphones. This process is described in the following.

As seen in Figure 7, the B&K response (dashed) has prominently higher level at low frequencies. The compensation of this large difference was not seen reasonable, since it would have increased the level of microphone self-noise significantly. Thus, the low frequency boosts in target responses were first flattened by applying a shelving filter [20]. Then the target responses were smoothed to

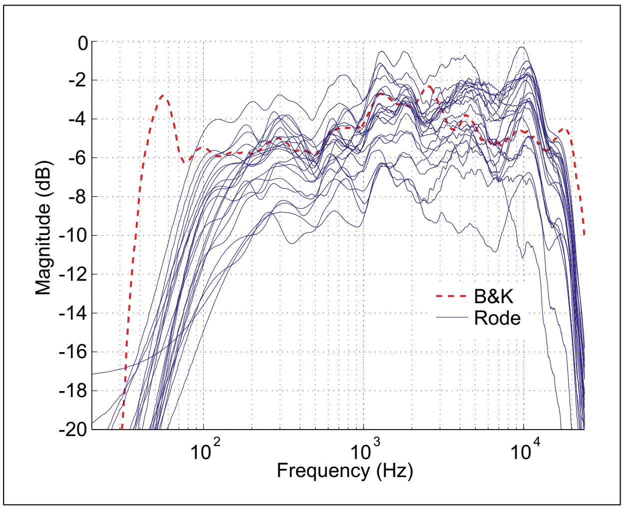


Figure 7. Original measured magnitude responses of each of the 22 microphone channels in their final positions. The dashed line is the response of the B&K microphone in the large anechoic room.

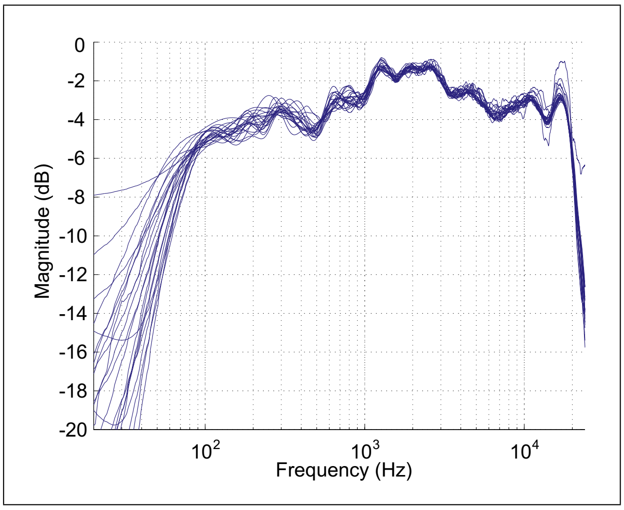


Figure 8. System corrected magnitude responses of all 22 microphone channels. Note that responses are not totally flat, since the loudspeaker response is included.

a 1/3 octave resolution. To give more weight to low frequencies in filter design, samples from target responses were chosen according to the ERB resolution [21]. Finally, IIR filters of order 50 were fitted to sampled targets with the `yulewalk.m` function in the Matlab environment. Despite the low frequency weighting, the designed filters did not perfectly match the targets at low frequencies. Therefore, another manually-tuned shelving filter was applied to each response in order to raise the low frequency response back to suitable level. As a result, the equalization of each microphone signal was performed with one IIR and one shelving filter. The corrected system responses are shown in Figure 8 where it can be seen that the responses are now within 1 dB range between 500 and 10 000 Hz and within 2 dB range from 60 to 20 000 Hz. It is notable that in these microphone responses the loudspeaker response is still present.

Possible sources for error in the measurements can be caused by the additional equipment present in the recording chamber. In spite of the apparent proximity of the extra loudspeakers (see Figures 3 and 5), the probability of sound energy reflecting to the microphones from the loudspeakers is diminutive. All rigid structures in the room supporting the equipment are padded with absorptive material. On the other hand, the absorbing wedges in the room can cause slight deviation between the microphone positions. Since the wall reflections are different depending on the amount of the wedge material in certain position, this can cause variation in microphone responses. However, the microphones were not moved from their final measurement positions.

4. Music material

Since the size of a typical orchestra and the complexity of the music texture varies between periods, recordings of different music styles can provide more information about the acoustics in auralization. Different acoustics are also often preferred in performance of different works. To have a selection of anechoic music excerpts suitable for room acoustical studies, passages representing Classical and Romantic styles were chosen for the recordings. While Baroque or Modern works have their own typical tone, works of these styles were not seen to offer considerably new information in acoustical sense. Modern and contemporary compositions have also copyright protections, which would prevent distributing the recordings freely. Additionally, Baroque orchestral works often include parts written for cembalo, which would have introduced further challenges considering the recording environment. In the following each recorded music excerpt is described briefly.

A soprano aria of *Donna Elvira* from the opera *Don Giovanni* by *W. A. Mozart* (1756-1791) was selected to represent typical music of the Classical period. This is the only recorded passage here which includes a soloist. On the other hand, the small orchestra characteristic to this era is the smallest of the recorded works. Besides the string instruments, parts are written for a flute, a clarinet, a bassoon and two French horns. The number of first and second violins in orchestra is typically around 10 players each. As this is the only piece recorded in whole length, it has a duration of 3 min 47 sec, thus being the longest of all four passages.

L. van Beethoven's (1770-1827) Symphony no. 7 was chosen due to the big chords and string *crescendos* in the introductory part. In auralization, the chords and pauses make the reverberation tail of a concert hall clearly audible, which was one of the reasons for selecting this work. For the musical style, it represents the late Classical period. The score includes parts for two flutes, oboes, clarinets, bassoons, French horns and trumpets in addition to the strings and timpani. The size of the string sections is slightly larger than in Mozart's music, as 12 first and second violins is a typical number. Bars 1-53 from I movement were recorded and the duration is 3 min 11 sec.

A. Bruckner's (1824-1896) Symphony no. 8 in turn represents the late Romantic period, and the overall dynamics of the music as well as the size of the orchestra are large. The score is written for a full symphony orchestra, containing parts for trombones as well as a tuba. Long passages for strings are written in *tremolo*, which is typical to all Bruckner's symphonies. While the texture is still quite conventional, it is noticeably dense, and many sections are played in *tutti* and *fortissimo*. The recorded section contains bars 1-61 from the II movement, thus being the shortest passage with duration of 1 min 27 sec.

G. Mahler's (1860-1911) Symphony no. 1 was selected as another late Romantic composition. As the music is composed in the same period as Bruckner's symphony, they are both great examples of works which require large orchestras. However, compared to Bruckner's music, the texture in Mahler's symphony is considerably more complex. Bars 1-85 from the IV movement were recorded. The recorded excerpt has a duration of 2 min 12 sec.

4.1. Conductor video

A video of a conductor was recorded with a digital video camera with external microphones. Only the conductor was included in the picture.

The pianist who played on the conductor video was a professional *répétiteur* but also a conductor as well. The piano track was played from the conductor's scores. Despite the complexity of the full score, particularly in Mahler's symphony, the pianist managed to play all essential details and followed the conductor well.

Mozart's aria was recorded with a soprano soloist in order to provide more predictable reference for the musicians playing along the video in recording situation. Finally, the four selected passages were edited to separate Quicktime movies.

5. Recordings

5.1. Musicians

Professional musicians were collected from the Finnish National Opera, the Finnish Radio Symphony Orchestra, the Helsinki Philharmonic Orchestra, and Tapiola Sinfonietta. The conductor on the video was not from any of these orchestras. Only one musician per instrument played all parts one after another. In total, 14 musicians were recorded, each session lasting from 1.5 to 6 hours. The largest number of different instruments recorded was 19 in Mahler's symphony. This number includes individual percussion instruments, a piccolo as well as two different clarinets and trombones. For Beethoven and Bruckner, 11 and 15 instruments were recorded, respectively. The lowest number of 9 instruments was required in Mozart's aria.

As large orchestras have up to 16 players per a violin part, a single violinist alone cannot produce similar breadth in the sound as several musicians playing in *unisono*. However, based on our previous studies [22] it should be enough for auralization purposes to record only one of each string instrument. Furthermore, several takes for each part were recorded. Therefore multiple different takes with only a little faults were used after all to produce the impression of several musicians.

One problem was predicted relating to the anechoic recording environment. Because the room did not provide any acoustic support, especially the string players were expected to push their playing in order to produce louder sound, which would be easily audible as bad sound quality. Therefore, each musician was specifically instructed not to use any excess force but instead to play similarly as in a performance situation.

Every musician was first introduced to the anechoic room and the recording system when arriving to the recording session and instructions on the correct positioning and playing style were given. Each instrument was tuned with an "a" note recorded from the same piano as in the conductor video (a = 442 Hz), although some musicians also used a tuning meter.

Actual recordings were usually commenced with Mozart or Beethoven, as these passages were the most conventional and gave a good opportunity to get familiar with the recording procedure. As expected, these excerpts were completed with least takes or challenges. On the other hand, they both contain very delicate sections requiring accurate and consistent timing.

Mahler's and Bruckner's symphonies required recording parts in shorter sections. Most of the brass instrument parts in Mahler's symphony were recorded in multiple segments, since many of the parts included sections where it seemed to be particularly easy to accidentally hit a wrong tone.

Recording some instruments in Bruckner's symphony was approached rather differently than written in the score. With violin parts, keeping in tune in long, high notes proved out to be very difficult but also exhausting while playing *tremolo* in *fortissimo*. Therefore an alternate method was applied. First acceptable takes of all *divisi* in the whole passage were played in steady sixteenth notes with correct dynamics. Thus, it was possible to play the excerpt entirely without fatigue and still keeping in tune. After that second versions were recorded in *tremolo* but this time softly in *mezzo-piano*, mostly disregarding the dynamic indications. Thus, we obtained separate takes for different playing techniques that were utilized in the post-processing later on. As viola parts were recorded shortly after violins, the method described above was utilized from the beginning with success.

Of all recorded instruments only one, tuba, succeeded in producing sound level that exceeded the maximum level in one microphone channel. During the recording of Bruckner's symphony, the signal of the microphone just above the bell was distorted. Fortunately, this was noticed immediately after the take, and a new one was recorded with slightly softer tone to prevent clipping.

Each player were given opportunities to listen to their recorded parts once in a while during breaks. For instance, the best takes from all French horn parts were combined so that any need for performing another take would have been immediately recognized.

After all other instruments, the soprano soloist was recorded. Because the original conductor video of Mozart's aria included the soprano, a second version of the video was constructed before the actual recording. Here the audio track was changed to the ensemble of instrument recordings that were already completed at that time.

5.2. Comments from musicians

All musicians were enthusiastic about this project and they were curious to hear the final results. Some of the players were interested in the directivity of their own instrument, although only preliminary results were visible by comparing the recorded signals from different microphones.

Some musicians were slightly wary of the recording environment, as this was their first visit to an anechoic chamber. Despite the unnatural environment, all musicians easily adapted to the situation and were able to play with good intonation and high quality.

Some comments were received concerning the style of the conductor. Since the musicians were from different orchestras, each player interpreted the conducting beat on the video by the tradition of the orchestra of the particular musician. However, the piano track helped to quickly find a similar interpretation.

6. Post-processing

To gather takes from all recorded instruments and to form an ensemble playing together, editing was required. After recording all the planned material, the prospective takes for each part were selected by listening carefully to all the accepted takes for finding missed notes or off-tempo passages.

In the first editing stage, one complete take was joined from several clips, if necessary. A common task was to replace accidental wrong notes on wind instruments in an otherwise good take. All editing was performed in sample-accurate manner, thus the length of resulting files were kept unchanged. At this stage any further editing was not performed.

Second editing stage comprised importing a whole instrument part to REAPER audio editing software. This software allowed the simultaneous editing of all 22 microphone tracks as this feature is very important for maintaining synchronization in all channels. This editing stage was essential for correcting any timing inaccuracy between the instruments. As a sophisticated feature in the multi-track software, crossfades were automatically created for all edit points, which provided satisfying results rather easily.

The editing process in the latter software was performed as follows. First imported parts in each passage were edited by using the piano track as a timing reference. These parts included usually some string instrument parts and a wind instrument. Timing inaccuracies were corrected, and finished stacks of 22 microphone tracks were rendered into individual files. After the first completed parts, the piano track was muted and the actual instrument recordings were used as timing reference from this point forward. One microphone channel for each of the completed parts was then left to represent the part in question for the editing of next imported parts. This cycle of importing, editing and rendering was repeated until all parts were completed.

Finally, separate files were processed in Matlab environment with the equalizing filters described in Section 3.4.

The goal in editing was not to create an unnaturally accurate synchronization. Therefore slight timing discrepancies were left unchanged. However, all the corrections were attempted to accomplish in a delicate manner so that the edits would not be easily perceived even by listening individual tracks.

Of the four recorded music examples, Mozart's aria was considered to require least editing as it had the smallest number of instruments. However, as the soloist was in the lead with the conductor on the reference track, the pianist had to adapt to the tempo more than in other pieces. This caused some irregularities in the rhythm and ultimately led to some passages being slightly out of tempo. The aria also featured a *fermata* pause in the middle. Synchronizing the tempo right after this pause required minor adjustments in all parts. Thus, it is noticeable that musical works allowing a soloist more freedom are very sensitive to the reference track.

On the other hand, Beethoven's Symphony no. 7 was regarded the most challenging from the synchronization point of view. This was noticed during editing, as more deviations concerning the rhythm had to be corrected in long sixteenth note scales and delicate segments requiring accurate articulation. Parts played by the same musician were better in time with each other than with other instrument parts. This indicates that a professional player can maintain similar interpretation through a recording session. The first chord did not require as much editing as anticipated, although minor adjustments were necessary.

Recording of Bruckner's symphony was not presenting any serious problems in editing. While it resulted in the largest number of tracks and parts, the rhythm in the texture is straightforward, thus being easier for the players to follow in tempo. Most editing work was caused by the inaccuracies between instrument groups.

Of the larger number of complete takes of violin and viola parts, more than just one usable take could be constructed. In Bruckner's symphony, a total of 20 string instrument tracks were edited. This is expected to be beneficial in the future applications by providing a richer sound, as consecutive takes are always a little different. Even more convincing imitation for an instrument section could have been achieved by recording multiple musicians with multiple instruments.

The parts in Mahler's symphony with more complex rhythms were comparably well in tempo. The introductory part presented some inaccuracies in the beginnings of long notes in *unisono* after rapid quintuplets and sixteenth note passages. Towards the end of the excerpt, the piano track had sudden changes in tempo. This was reflected to the violin part recordings and is noticeable even after editing attempts.

The number of performed edits was approximately 2-3 on average in each accepted take. The piano track on the conductor video was regarded very helpful in maintaining synchronization and keeping in tune. Fortunately, any major audible problems with playing in tune were not experienced, as corrections would be nearly impossible in the post-processing stage. The need for editing would have most certainly been enormous without the piano track.

Interestingly, hearing the piano while playing had a strong influence on the rhythm of actual playing. This was noticeable in some passages in which the piano was not exactly in tempo. As a result most players followed the piano instead of the conductor. However, individual timing mistakes in the piano track did not affect to the playing in tempo.

Another result was noticed while editing different parts. In some passages, all parts performed by a single musician featured a common dissimilarity compared to parts on other instruments. Therefore a few notes needing editing on one part often indicated an upcoming need for editing on other parts as well.

6.1. Noise analysis

Noise levels were not accurately measured during the system set-up, although some test recordings were made with a clarinet to assess the overall sound quality. Spectral analysis of recorded sections with silence reveal that most of the noise is at very low frequencies.

Noise in a single channel is not noticeable, but with multiple parts mixed together the noise level becomes higher. As an example we measured A- and linear-weighted noise levels in 47 combined instrument tracks of Bruckner's symphony. The measured levels are shown in Table III. For a reference, peak level of the combined tracks was at -6.3 dB resulting in a dynamic range of 44.5 dB.

During pauses instrument channels contain only noise. To reduce overall noise in mixed recordings the use of noise gate function built-in to the editing software was investigated. This functions as an automated attenuator to mute a channel if the signal level remains under certain threshold. Adjusting the sensitivity is important in two matters. First, the gate have to open well before the actual instrument sound is present. Second, the gate has also to close slowly enough to prevent audible change in overall noise level. However, very moderate noise levels of approximately -70 dB do not present problems in finding suitable parameters. Values presented in Table III are obtained with discrete gate parameters. As the gated noise values are much better allowing 57.2 dB signal to noise ratio in Bruckner's symphony, this method can be potentially used in future studies.

Only a limited number of takes were recorded of the string parts. To reach correct balance between the instrument groups, these tracks have to be amplified and multiplied, which increases the signal as well as noise level in strings recordings. Hence, the noise analysis and gating is important for increasing the quality in auralization.

Table III. Measured noise levels with and without A-weighting in recordings of Bruckner's symphony. Flute track is from the microphone in front direction. I: 1 track (Flute 1); II: 47 tracks mixed.

	original		gated	
	$L_{\mathrm{noise,A}}$	L_{noise}	$L_{\mathrm{noise,A}}$	L_{noise}
I	-77.9dB	-71.2dB	–	–
II	-57.3dB	-50.8dB	-70.0dB	-63.5dB

7. Future work

The motivation for the presented recordings was twofold, to produce high quality stimulus for auralization studies and to gather directivity data of musical instruments.

In auralization studies the performed recordings can be applied at least in two ways. First one is to take one single microphone recordings of all instruments and apply these in multiple point sources in a concert hall model. The directivity of the instruments can be taken into account by filtering this single signal according to the directivity of the particular instrument, as presented by Savioja *et al.* [23]. Another way to represent a sound source in auralization is so called multi-channel auralization, in which a point source emits different anechoic signals to different directions [12].

Another usage of the anechoic symphonic recordings would be in evaluations of concert hall acoustics. The individual instrument recordings can be played from dozens of loudspeakers distributed on the stage of an existing concert hall. This "loudspeaker orchestra" can be listened in-situ in the hall, but it can also be recorded, e.g., with a binaural artificial head. When the whole reproduction and recording system is calibrated carefully, such a loudspeaker orchestra enables an A/B test of different concert halls in the laboratory environment. Although the directivities of the loudspeakers are not exactly the same as directivities of musical instruments, this method looks very promising in the future studies of concert hall acoustics.

8. External resources

Two-channel downmixes of anechoic music and example concert hall auralizations are available in the Internet. In addition, all the individual audio tracks are freely available for academic use on request. See more information at http://auralization.tkk.fi.

9. Conclusions

The anechoic recording of symphony orchestra instrument by instrument was implemented with a dodecahedral microphone setup. As a result, multichannel anechoic symphony orchestra recordings were obtained from four different musical styles ranging from Mozart to Mahler. Experiences of the recordings and post-processing are discussed from different point of views. Recorded material are planned to be used in concert hall auralization studies as well as in investigating existing halls with repeatable excitation signals. The result of the project, four passages of symphonic music are provided to be used freely for academic research.

Acknowledgement

This work was supported by the Academy of Finland (projects nos. 105780 and 119092).

References

[1] M. Kleiner, B.-I. Dalenbäck, U. Svensson: Auralization – an overview. J. Audio Eng. Soc. **41** (Nov. 1993) 861–875.

[2] A. Krokstad, S. Strom, S. Sørsdal: Calculating the acoustical room response by the use of a ray tracing technique. J. Sound Vib. **8** (1968) 118–125.

[3] M. Vorländer: International round robin on room acoustical computer simulations. Proc. 15th Int. Congr. Acoust. (ICA'95), Trondheim, Norway, June 1995, 689–692.

[4] U. Svensson, U. Kristiansen: Computational modeling and simulation of acoustic spaces. AES 22nd Int. Conf. on Virtual, Synthetic and Entertainment Audio, Espoo, Finland, June 15-17 2002, 11–30.

[5] S. Siltanen, T. Lokki, S. Kiminki, L. Savioja: The room acoustic rendering equation. J. Acoust. Soc. Am. **122** (September 2007) 1624–1635.

[6] J. Blauert: Spatial hearing. The psychophysics of human sound localization. 2nd ed. MIT Press, Cambridge, MA, 1997, 494.

[7] H. Møller, M. S. D. Hammershøi, C. Jensen: Head-related transfer functions of human subjects. J. Audio Eng. Soc. **43** (1995) 300–321.

[8] Bang & Olufsen: Music from Archimedes. Audio CD 101, 1992.

[9] Denon: Anechoic orchestral music recording. Audio CD, Denon Records, 1995. ASIN: B0000034M9.

[10] T. Hidaka, K. Kageyama, S. Masuda: Recording of anechoic orchestral music and measurement of its physical characteristics based on the auto-correlation function. ACUSTICA **67** (1988) 68–70.

[11] R. Freiheit: Historic recording gives choir "alien" feeling: In anechoic space, no one can hear you sing. 150th ASA meeting / NOISE-CON 2005, Minneapolis, USA, Oct. 2005. paper 4pMU3.

[12] F. Otondo, J. H. Rindel: A new method for the radiation representation of musical instruments in auralizations. Acta Acustica united with Acustica **91** (2005) 902–906.

[13] M. Vigeant, L. Wang, J. Rindel: Investigations of multichannel auralization technique for solo instruments and orchestra. 19th International Congress on Acoustics (ICA'2007), Madrid, Spain, September 2-7 2007. Paper RBA-15-004.

[14] J. Meyer: Acoustics and the performance of music. Verlag das Musikinstrument, Frankfurt/Main, 1978.

[15] J. Meyer: The sound of the orchestra. J. Audio Eng. Soc. **41** (Apr. 1993) 203–213.

[16] Directivities of musical instruments. http://www.ptb.de/en/org/1/17/173/richtchar.htm.

[17] N. H. Fletcher, T. D. Rossing: The physics of musical instruments. Springer, New York, 1991.

[18] Genelec 1032A datasheet. http://www.genelec.com/documents/datasheets/DS1032a.pdf.

[19] S. Müller, P. Massarani: Transfer function measurement with sweeps. J. Audio Eng. Soc. **49** (June 2001) 443–471.

[20] U. Zölzer (ed.): DAFX – digital audio effects. John Wiley & Sons, May 2002.

[21] B. Moore, R. Peters, B. Glasberg: Auditory filter shapes at low center frequencies. J. Acoust. Soc. Am. **88** (July 1990) 132–140.

[22] T. Lokki: How many point sources is needed to represent strings in auralization? International Symposium on Room Acoustics (ISRA'2007), Seville, Spain, September 10-12 2007. Paper P11.

[23] L. Savioja, J. Huopaniemi, T. Lokki, R. Väänänen: Creating interactive virtual acoustic environments. J. Audio Eng. Soc. **47** (1999) 675–705.