Field Report: A pop production in Ambisonics

LAC 2010, Utrecht

Jörn NETTINGSMEIER
Freelance audio engineer and qualified event technician
Lortzingstr. 11
Essen, Germany, 45128
nettings@stackingdwarves.net

Abstract

This paper describes an acoustic pop production in mixed-order Ambisonics, using an ambient sound field recording augmented with spot microphones panned in third order.

After a brief introduction to the hard- and software toolchain, a number of miking and blending techniques will be discussed, geared towards the capturing (or faking of) natural ambience and good imaging. I will then describe some peculiarities of Ambisonic mixing and the struggle to make the resulting mix loud enough for commercial use while retaining a natural and pleasant sound stage and as much dynamics as possible.

Keywords

Ambisonics, surround sound, pop music production, mixing techniques

Introduction

In Feb 2010, German singer/songwriter Tom Gavron scheduled a recording session featuring three different line-ups: a quartet featuring piano, violin, cello and percussion, a duo with piano and bassoon, and a jazz sextet. Overdubs were to be limited to the vocal tracks, to capture a natural group feel and allow for improvised interaction.

With so many interesting acoustic instruments, it became clear that their spatial characteristics and interaction with the room ambience had to be captured, rather than relying on panned mono sources as usual.

Although I had no idea of the recording venue acoustics, I decided to try an Ambisonic approach, using a tetrahedral main microphone backed up with a standard close-miking setup.

Since the recording was to serve both as promotional material and merchandise, it was clear that an easily accessible and distributable stereo mixdown was the primary target. In addition, we planned to create a 5.1 mix with an "on-stage" perspective as an extra feature for fans with the necessary equipment, plus a native B-format release for the limited number of Ambisonics enthusiasts out there.

Equipment

The session was recorded on a Lenovo ThinkPad X61 running a heavily customized openSUSE 11.1 with kernel 2.6.31.12-rt20 and current SVN heads of JACK (r3898), FFADO (r1794) and Ardour2 (r6635).

The audio data was written to an external USB 2.0 drive[1] and backed up to a second harddisk every night.

A Focusrite Saffire Pro 26 served as the recording interface. Its eight microphone preamps were complemented with another eight from an RME Micstasy, and eight cheap Behringer ADA 8000 channels for line signals and less important microphones. All units were connected via ADAT and externally synced to the Saffire's wordclock out.

---

[1] I found external disks to be more dependable than built-in notebook drives, since they have less tendency to overheat, deliver better sustained write rates and generally run more quietly. As an additional bonus, they can help circumvent shared-interrupt problems with the built-in controller and other important parts of the signal chain (as was necessary here since the SATA chip shares an IRQ with the cardbus controller).

A Core Sound TetraMic (see appendix) was used as the main microphone.

On the software side, Ardour2 [Dav10] was used both for tracking and mixing.[2]

Recording

Using a "main microphone" approach in pop music may seem strange, and it does have a few pitfalls. The soundstage is more or less determined during setup. While you can drag instruments away from their natural location with the spot mikes, you will risk incongruent directional cues if you overdo it. From early on, you have to get your client to cooperate and refrain from fancy panning suggestions in post-production. The prize will be a beautiful, natural ambience.

1   Quartet session

The first two studio days were dedicated to the violin/cello/percussion/piano line-up, and were recorded in a rather small room (about 4x7m) that was acoustically treated as a percussionist's rehearsal room, i.e. very dry but pleasant.

With no separate control room available, the recording gear (and the engineer) ended up in the same room, which greatly eased the communication with the performers. The mike setup however was rather tedious, since several test recordings of every instrument were required to be able to judge the recorded sound accurately.

The four musicians were arranged in a circle, facing each other, and the TetraMic was placed in the centre, slightly favoring the strings to ensure a good natural balance.

Spot microphones were used as follows (from rear left to rear right):

On the **pedal timpano**, a trusty old AKG D-112 about 10cms away from the head, close to the rim.

The **tom-tom** was handled by a Sennheiser e904.

A cheap Sennheiser vocal mike that we had lying around was put on the **snare** after it was found that a Røde NT5 could not cope with the sound pressure up close, even with the comparably light touch of a classical drummer.

---

[2]Wiring Ardour for Ambisonics is outside the scope of this paper. See [Net09] and [Net09-2] for a detailed explanation.

Illustration 1: Quartet session. Interesting mikes, top to bottom, left to right: coincident pair of AKG CK91s for drum set; Core Sound TetraMic as main microphone, offset towards the strings; AKG CK91 violin spot; BPM CR-95 cardioid cello spot

The entire set (which also included a number of splash and ride cymbals and a hi-hat) was covered with an **overhead** XY pair of AKG CK-91 cardioids at a height of about 2 metres, pointing almost straight down.

The **electric piano** (a Korg StageVintage) has very nice balanced outputs which were used to capture the direct signal. Its stereo jack outs were routed to two active RCF cabinets placed behind the pianist on the floor, tilted upwards like a monitor wedge.

This way, the electric instrument blended nicely with the room and the acoustic instruments, and the TetraMic could capture some meaningful directional information.

For the **violin**, we chose another AKG CK-91 and placed it 20-30 cm above the upper part of the fingerboard, pointing at the bridge.

The **cello** was covered with a BPM CR-95 switchable large dual-diaphragm in cardioid setting, 30cm away from an f-hole slightly below the bridge. It was shielded from the drums with some jackets draped over a music stand (not shown in photo), to reduce crosstalk from the snare drum.

All musicians were facing each other, and the nulls of all microphones were kept pointing inwards as much as possible, to improve channel separation.

For the entire duration of the session, all tracks were kept armed and recording, with short breaks

every hour or so, to save clean snapshots while the transport was stopped.[3]

Since the material to be recorded was new to the musicians, the arrangements evolved considerably during the session. It also meant that preciously few compatible parts were available for editing.

In the end, most of the recordings were entire takes (usually between 13 and 20 per day), the last few of which where usually selected for post-production with minor edits to be made. Some solo parts were recorded as separate takes to reduce stress and fatigue, but always "live", i.e. with the entire band.

A click was used to ensure a consistent base tempo before each take, but the recordings themselves were done without.

The vocals were to be overdubbed on a separate date, allowing time for careful selection of the final material and preliminary editing.

2   Piano/bassoon session

The room was 5 by 6 metres with a height of around 3 metres, and acoustically treated.

We were lucky to be able to use a Steinway **grand piano**, which was recorded with an ORTF pair of AKG CK-91s. The lid was propped up in low position, and the mikes were located to the rear of the instrument, slightly above lid level, to avoid the boomy quality of the direct sound coming through the opening.

Another CK-91 covered the **bassoon**. During warm-up, I learned that the sound of a bassoon travels along the entire length of the instrument, depending on register: the low notes originate from the top of the bell (the upper part), whereas the high notes emerge from the boot (the lower part), with many positions in-between. To capture this interesting quality, the spot mike was augmented to an M/S pair with the BPM CR-95 in figure-of-eight setting.

Additionally, a Røde NT-5 was aimed at the bell from above, to have some additional wind noise and more pronounced overtones for flexibility in the mix.

All instruments and microphones were lined up along a common axis, the microphones pointing outward to reduce crosstalk. The main microphone was placed well outside the center of the room at a height of about 2m, so that the instruments

subtended an angle of around 120 degrees. The resulting hole in the soundstage was reserved for the vocal overdub.

3   Sextet session

The jazz sextet consisted of a standard jazz drum kit, double bass, piano/keyboard, guitar, trumpet/flugelhorn and vocals. Again, the session was captured with a TetraMic in the centre. This time, the spots were used rock'n'roll fashion, i.e. extremely close, to get some channel separation.

The **drums** were covered with three Beyer clip-ons for snare and toms, an AKG D-112 for the kick, and two AKG CK-91 in XY configuration for overheads.

The **double bass** was recorded to two tracks: a DI signal from a piezo pickup and the BPM CR-95 in cardioid setting, positioned very close to an f-hole.

Regrettably, no guitar amp was available for the session, so the **guitarist** plugged his ES-175 and a Telecaster into a Lexicon MPX-100 which was driving a small active RCF P.A. cabinet miked with a good ol' SM57. Needless to say, the attack was shoddy.

The Steinway **grand** had to be kept closed so as not to upset the natural balance between the instruments in the room. Hence, the mike (an ORTF pair made of Røde NT5s) had to be placed at the only available opening, about 30cms above the tuning pegs. The coverage of the instrument in the extreme registers was not too good, but for the reduced jazzy playing style, which mostly featured the middle register, the compromise was acceptable.

The **trumpet and flügelhorn** were captured with another CK-91 with 10dB pad, aimed well above the bell. The player was instructed to lift the instrument slightly for emphasis, so that extra brilliance would be captured whenever he felt necessary.

The **vocalist** used a hand-held Beyer MCE 91 for a live take of Sinatra classic "Come fly with me".

The second piece to be recorded was an instrumental rendition of "My funny valentine", for which the vocals were to be overdubbed later.

As expected, snare and cymbals leaked into most microphones. For a few dBs of crosstalk reduction (which can make a world of difference in the mix), an empty gear case was used as a barrier between

---

[3] This precaution turned out to be unnecessary, since Ardour saved reliably even while transport was rolling.

piano mike and drums, and a couple of winter coats hung over a mike stand served as a shield for the bass mike.

4 Vocal overdubs

The vocal dubs were done two days after the main session, in a 4x5m living room with wooden flooring and one wall deadened with a large mattress.

We used the BPM CR-95 in cardioid setting, with a windscreen in front, positioned slightly below mouth height to ensure a relaxed singing posture and avoid the tendency to lift the chin while singing.

About half a metre behind the main mike, the TetraMic was running along for some optional room ambience.

The singer was fed a mixture of the main mike and a virtual Blumlein array from the TetraMic. It was pointing upwards, away from the direct sound as much as possible, to avoid coloration. I find that a good room signal helps reduce the listening fatigue induced by closed studio headphones. Since it feels natural even at low levels, it does not affect intonation as much as artificial reverb, which usually has to be turned way up to for a comfortable listening experience.

A stereo fold-down of the TetraMic recording of the basic tracks was used as the primary monitor signal, complemented with some dry piano for pitch reference, and additional direct signals as required by the vocalist.

The overdub takes consisted of one four-channel track for ambience and one mono track.[4] They were recorded on top of one another during the session.

To sort the material, we created four new pairs of tracks, to which the recorded takes were moved after listening: one for trashcan, one "maybe usable", one "satisfactory" and one for material deemed very good. Ardour's "Lock edit" mode proved very helpful, as it eliminates time alignment errors when dragging material between tracks. Good takes with questionable parts in them were split to exclude the problematic section.

---

[4]In retrospect, it would have been better to record the vocal takes to a single five-channel track each, to avoid confusion in the editing stage. Such a compound track is easily split into manageable parts using Ardour's flexible buses.

5 Post-production

5.1 Editing

Edits were done in the recording configuration, using a standard stereo master bus, and monitored through stereo speakers. Each edit was then cross-checked with headphones.

As expected, the main microphone technique and associated cross-talk made convincing edits very difficult, and the assembly process was time-consuming. After some experimenting, it was found that staggering the cuts of main mike and spots helped hiding otherwise questionable edits: minor problems with overhanging sounds and long crossfades in the ambience would become acceptable if a clean note onset had been established in a spot mike before.

Again, the "slide edit" and "lock edit" modes of Ardour were used heavily. "Slide" allows regions to be dragged around in time, and is needed to align an insert with the groove. When that has been done, "lock" fixes all regions in time and only permits the trimming of region boundaries – that way, the edit can be cleaned and made inaudible, without the danger of messing with the groove unintentionally (which happens rather easily with Ardour when screen space is limited and track heights are small).

6 Mixing

For mixdown, a new 16-channel summing bus was added for third-order Ambisonic mixdown, and the old "master" bus was deleted. Additional two-, four-, and nine-channel monitoring busses were created for monitoring in UHJ-encoded stereo, first and second order Ambisonics.

6.1 Using convolution reverb

Spot mikes need some additional reverb to sound natural at the listening spot defined by the position of the main microphone. Ideally, this reverb is an impulse response of the recording room recorded by the main mike, where the excitation speaker is placed at the location of each spot microphone. For maximum fidelity, separate IRs should be captured for each instrument group (or even every microphone position).[5]

---

[5]Aliki [Adr09] is a good tool for the job. The capturing of room responses is described in detail in the Aliki manual.

These IRs can then be combined into a convolution matrix for an engine such as jconvolver[6], so that it has N inputs, one for each of the IRs, and either two or four outputs, depending on whether the room response was recorded in stereo or first-order B-format.

Optionally, the early reflections and tail section of an IR can be separated. One tail can then be used globally (because the tail does not contain significant directional cues and differences from one position to another are very subtle at best), and only the short early reflection parts are treated individually. This conserves CPU and allows for an extra degree of flexibility, namely the ratio of early reflections to reverb tail.

To plug such a beast into ardour, create an N-channel bus with a corresponding N channel insert connected to the external jconvolver. Only the first few return channels will be used, and the rest can be left unconnected. Similarly, only the four active outs of the N-channel bus will be connected to the first four channels of the master bus which contain the zeroth- and first-order components.

Unfortunately, no adequate speaker for an IR measurement of the recording rooms was available, so some "foreign" IRs had to be used on the spot mikes to blend them with the slightly wetter room signal.

6.2   Source alignment

The natural sound stage as recorded by the TetraMic was used as the basis for source positioning. With stereo in mind, the sound field was rotated to provide a not-too-unconventional balance when folded down, and to leave space for the singer in the front. Where possible, strings were placed in the back, since they benefit from the slightly phasy, blurry quality which UHJ stereo encoding adds to rear sources. For the same reasons, bass instruments should be placed in the front quadrant if possible.

Spot mikes were brought up one by one and aligned with the Tetramic sound stage by ear, moving them until the sources stopped "jumping" when switched.

Each spot mike was delayed to compensate for the offset to the main microphone, to avoid comb filtering effects in the combined signal. Sometimes, the actual delay used differed from the measured value by several milliseconds, if a more pleasant timbre could be obtained.

6.3   Equalisation

As shown in the photo, the recording situation was rather cramped, resulting in a pronounced bass boost in the microphone due to proximity effect. When played back, the sonic impression was quite obtrusive, although technically correct. Some bass reduction in combination with gentle reverb was employed to give the mix a more spacious feel and make the instruments back away from the listener.

Signal crosstalk was quite bad, and steep high-pass filters had to be used on almost every microphone to keep the timpani and bass drum out.

The extreme close-miking of the strings produced a slightly harsh tone that proved difficult to correct without losing the "shimmer" of the bow sound. In the end, reverb was more effective than filters.

During post-production, it was found that an EQ setting that works for the UHJ-encoded signal will also sound good in Ambisonics, but not vice-versa. The full Ambisonic rendering is a lot more spacious and transparent, and can absorb more reverb. One must resist the urge to "fatten it up" too much, as this will result in a boomy and overly wet UHJ stereo image.

Frequently, instruments which had been very pleasant when soloed sounded tinny or otherwise artifical in the mix. This is a common phenomenon when many microphones are open, and there is no way around it other than to paint the spot mike sound "with the large brush" (i.e. to over-exaggerate the desired characteristics a little), to close unused microphones wherever possible, and to keep fiddling with the delays. In retrospect, hypercardioid patterns wounld have helped to reduce crosstalk between adjacent instruments.

6.4   Building the mix

My usual approach is to start with the drum overheads and any room microphones, add spots one by one, assemble a basic track, and bring the vocals in at the end. Since the Ambi setup reacts very differently than a standard stereo system, I found that I had trouble finding room for the vocals this way and gave up after a few failed

---

[6][Adr10]; jconvolver is unique among the freely available JACK convolvers in that it uses a variable partition size and can be configured to incur only one period of latency regardless of IR length.

attempts. Instead, I started with the vocals and a little bit of piano, making sure the song would work as-is. The other instruments were then tucked "under" this basic mix. Afterwards, the room microphone was brought up for some "air". Additional reverb was added to each signal individually, and finally, fader automation was used to emphasize dynamics, clean up the mix and add some final polish.

6.5   Dynamics

In the final stages of the ambisonic mixdown, it became apparent that commercial impact would require some sort of peak limiting and gentle overall compression. Regrettably, no free multichannel-capable compression tools are available at this point[7] (and Ardour2 cannot make use of a plugin's side chain port easily), so the master was left unprocessed and the individual channels were treated instead.

This is a serious drawback, since it entangles the mixing and mastering stages. Keeping them separate (and bringing in a fresh pair of ears) has its advantages: the mixing engineer does not have to deal with real-world playback systems and can create an artistic mix under optimum circumstances, and the mastering engineer can then deal with the necessary compromises to make it work on Joe Sixpack's car stereo. It takes some effort to focus on mastering processing and not constantly question and revisit earlier mixing decisions.

On the up side, the achieved mix was a lot more transparent than with sum compression, while the loudness was slightly lower than average. To compensate, some additional limiting was performed on the UHJ-encoded stereo output.

The automatic 5.1 folddown was done with Fons Adriaensen's hand-optimized second-order ITU decoder that comes with AmbDec as a default preset.

---

7 JAMin [Har05] or the very promising Calf multiband compressor [Fol10] come to mind. As it stands, even something as kludgy as a compile-time switch to allow for multichannel work with hardwired side-chaining would be highly welcome.

7   Other production approaches

Recording more-or-less "live", without click, is nice but not always practical. Far more material needs to be discarded for mistakes, there are less options for repair and improvement, and the overall level of perfection that can be achieved is limited unless the musicians are of the very best.

But you can easily use traditional single-instrument overdubbing in an Ambisonic production. Naturral room ambience (if desired at all) can either be faked using B-Format impulse responses as described earlier, or you can keep the main microphone running each time a musician lays down a track. In theory, the result will be the same as if everybody had played in the room at the same time. In practice, you will also get a lot more hiss. But to blend one or two soloists into the mix, this approach is feasible.

Conclusion

Doing a pop production in Ambisonics is definitely possible. One must constantly double-check the mix in both UHJ and Ambisonic renderings, but that is a fair deal compared to the hassle of an extra surround mixing session. With the available free software tools, most recording problems can be dealt with, and Ambisonic panning opens up new creative possibilities. Even on plain stereo systems, the sound stage can be extended well outside the usual stereo triangle, without complicated manual phase trickery.

A flexible multichannel compression tool with appropriate side-chaining would help get the job done more quickly. However, with a periphonic sound stage encompassing the entire sphere, sum compression is even more questionable than for stereo (where likewise the current best practice backs away from global dynamic processing and moves towards stem-based separation mastering).

It will be interesting to take this production approach to other genres, such as electronic dance music (where a slim chance of native ambisonic playback might exist in some clubs).

Appendix: The TetraMic

An implementation of a microphone design devised by Gerzon, Craven et al. in the 1970s, the TetraMic consists of four cardioid capsules arranged in the edges of a tetrahedron. Its native signal set (called A-format) can be converted into

the B-format used in Ambisonics by a simple matrix operation:

```
W' =  LFU + RFD + LBD + RBU
X' =  LFU + RFD - LBD - RBU
Y' =  LFU - RFD + LBD - RBU
Z' =  LFU - RFD - LBD + RBU
```

(where L/R means left/right, F/B is front/back, and U/D is up/down, to uniquely identify each capsule)

The signals are primed to indicate that some EQ correction is still missing to compensate for the slight positional error of the capsules (for a perfect microphone, they should be precisely coincident). For an easy-to-understand discussion of A-to-B format conversion, see [Far06].

On Linux systems, the conversion is handled by TetraProc [Adr09-2], whose author will provide a custom configuration file in cooperation with the microphone manufacturer.

The B-format can then be used natively for Ambisonic playback (either horizonal-only or full 3D), or an arbitrary number of first-order microphone patterns can be derived from it. In practice, one would create a coincident stereo pair, or a set of five (hyper-)cardioids for Dolby Surround. The big advantage is that orientation, opening angle(s) and polar characteristics can be selected during post-production, making it a very versatile main microphone.

In terms of localisation precision, the Tetramic is one of the best microphones of its design available today, owing to its small capsules which make the array nearly coincident to begin with, and the theoretically perfect digital post-matrix filtering. Its one great disadvantage is the low signal-to-noise ratio (a consequence of the small, cheap capsules), which makes it less well suited to very soft music such as a single acoustic guitar at a distance. For the task at hand, however, it was ideal.

The Tetramic is available from Core Sound LLC, http://core-sound.com. Quieter (and more costly) variants of the design are offered by SoundField, http://soundfield.com.

References

[Adr09] Fons Adriaensen, Aliki, http://www.kokkinizita.net/linuxaudio/downloads/

[Adr09-2] Fons Adriaensen, TetraProc tetrahedral microphone processor, l.c.

[Adr10] Fons Adriaensen, jconvolver, l.c.

[Dav10] Paul Davis et al., Ardour digital audio workstation, http://ardour.org

[Far06] Angelo Farina, A-format to B-format conversion, http://pcfarina.eng.unipr.it/Public/B-format/A2B-conversion/A2B.htm

[Fol10] Krzysztof Foltman, Thor Harald Johansen et al, Calf audio plugin pack, multiband compressor contributed by Markus Schmidt, http://calf.sourceforge.net

[Har05] Steve Harris et al., JAMin, the JACK Audio Mastering interface, http://jamin.sourceforge.net

[Net09] Jörn Nettingsmeier, Ardour and Ambisonics, in: eContact! 11.3 (Journal of the Canadian Electroacoustic Community), http://cec.concordia.ca/econtact/11_3/nettingsmeier_ambisonics.html

[Net09-2] Jörn Nettingsmeier, Using Ambisonics as production format, in: Beta Ardour reference manual, http://ardour.stackingdwarves.net/ardour-en/8-ARDOUR/279-ARDOUR/280-ARDOUR.html