Abstract
In this paper we present a library for 3D Higher Order Ambisonics (HOA) for the SuperCollider (SC) sound programming environment. The library contains plugins for all standard operations in a typical Ambisonics signal-flow: encoding, transforming and decoding up to the 5th order. Carefully designed PseudoUgens are the interface to those plugins to aim for the best possible code flexibility and code reusability. As a key feature, the implementation is designed to handle the higher order B-format as a channel array and to obey the channel expansion paradigm in order to take advantage of the powerful scripting possibilities of SC. The design of the library and its components is described in details. Moreover, some examples are given for how to built flexible HOA processing chains with the use of node proxies.

Keywords
SuperCollider, Higher Order Ambisonics.

1 Introduction
Ambisonics, i.e. the description of sound pressure fields through spherical harmonics decomposition, has been around for quite a while since its invention by [Gerzon, 1973]. Back in the days, the harmonics decomposition was up to the first order (First Order Ambisonics, FOA), using the 4-channel B-format. The playback of FOA audio content depended on special hardware and did not make it into mainstream audio in the first decade of its existence for various reasons, one of which being that FOA offers only limited spatial resolution.

Ambisonics research made significant advances in the 2000’s through the work of Bamford [Bamford, 1995], Malham [Malham, 1999] and Daniel [Daniel, 2000], who extended the sound pressure field decomposition up to higher orders hence the term (HOA). HOA increases the spatial resolution and thereby reduces the limitation of low spatial definition when compared with other spatialization techniques.

For streamlining and standardising content production, one hurdle that HOA was facing in the past was the coexistence of various channel ordering and normalization conventions. In order to address this issue, the Ambix standard was proposed by [Nachbar et al., 2011] and is ever since increasingly adopted by recent HOA implementations.

Today, processors can handle with ease multiple instances of multi-channel sound processes. Further, the rise of video games and Virtual Reality (VR) applications has elicited new interest in Ambisonics amongst audio researchers and content creators. This is mostly due to its inherent property to yield easy-to-manipulate isotropic 360 degree sound pressure fields, which can be rendered either through multi loudspeaker arrays or headphones. In the case of VR applications head-tracking is already available and the listener is always in the sweet spot. For the capturing of HOA 3D sound pressure fields, various microphone array prototypes have been developed some of them being available as commercial products like mhacoustic’s Eigenmike® for instance. As far as multi loudspeaker reproduction is concerned, the number of loudspeaker domes with semi spherical configurations is growing and electroacoustic composers have also shown increasing interest in HOA as a spatialisation technique, notably amongst them for composition and [Barrett, 2010] and sonification [Barrett, 2016].

1.1 Ambisonics in various platforms
Over the last few years HOA has seen various implementations in diverse sound software environments, mostly as plugins in DAWs. The Ambisonics Studio plugins by Daniel Courville, for instance, have been around for some time

http://www.radio.uqam.ca/ambisonic
Another recent and very comprehensive example is the Ambix plugin suite from [Kronlachner, 2013] also see [Kronlachner, 2014a]. For Pure Data and MaxMSP, HOA libraries have been made available by the Centre de recherche Informatique et Création Musicale [3] an early implementation can also be found with the ICST Ambisonic tools [Schacher, 2010]. An early version for Pure Data can be found in the collection of abstractions called CubeMixer by [Musil et al., 2003]. Recently, the HOA library Ambitools [4] developed mainly in Faust has been made available [Lecomte and Gauthier, 2015].

1.2 SuperCollider

The audio synthesis environment SuperCollider (SC) by [McCartney, 2002] is particularly well suited for the creation of dynamic audio scenes. SC is split into two parts: The server scsynth for efficient sound synthesis and sclang, an object oriented programming language for the flexible configuration and re-patching of DSP trees on the server. Similar to most sound programming environments, synthesis is based in SC on unit generators called Ugens. Third party Ugens are collected separately in SC3plugins. Extensions to scsynth are managed through Quarks. Ugens can be assembled to more complex arrangements through synthesis definitions, known as SynthDefs, which are executable binaries for synthesis in scsynth. In sclang, PseudoUgens can be created, which is another way of handling complex arrangements of Ugens in sclang, which are compiled for scsynth, when needed. For a detailed introduction to SC see [Valle, 2016] and the SuperCollider book [5] by [Wilson et al., 2011].

1.3 Ambisonics in SuperCollider

In 2005, Frauenberger et al. implemented HOA in SC as the AmbIEM Quark [6]. This implementation goes up to the 3rd order, and follows the old Furse Malham channel ordering and normalization. All unit generators (Ugens) like encoding, rotation, and simple decoding are implemented in sclang as PseudoUgens. AmbIEM comes with an simulation of early reflections in a virtual room but lacks functionality such as beamforming. The Ambisonics Toolkit (Atk) for SC by [Anderson and Parmenter, 2012] is

1.4 Library design in SuperCollider

In this context the paper presents a modern HOA implementation for SC, which is modular and adopts all established standards in terms of channel ordering and normalizations. Inspired by the approach found in the Atk and typical for the general design of SC, computationally intensive parts like Ugens are split from PseudoUgens convenience wrapper classes of sclang. The HOA library hence in three parts, SC3plugins, PseudoUgens and audiostream plus HRTFs for binaural rendering in a support directory.

1.4.1 SC3plugins

The first part of the library is a collection of Ugens, which is part of the SC3plugins collection. Each Ugen is compiled from C++ code. It consists of a SC language side representation of the Ugen as a .sc class file and a .scx, .so or .dll compiled dynamic link library, for the platforms (OSX, Linux or Windows) respectively. For each Ambisonics order (so far up to order 5), there are individual Ugens for the encoding, transforming and decoding processes in a typical Ambisonics signal flow. The C++ code for these Ugens is generated from the HOA library Ambitools [Lecomte and Gauthier, 2015] with the compilation tool faust2supercollider. This approach was taken for two reasons:

First, to leverage the work already accomplished in Faust. Indeed, the Faust compiler generates very efficient DSP code and the Faust code base allows to efficiently combine existing functionality. The meta approach through Faust will lead to future additions of functionality, which can then be easily integrated in the HOA library for SC.

Second, each Ambisonics order comes with a defined multichannel B-format, this in turn defines the amount of input and output arguments for the Ugens. For instance, a Ugen rotating an Ambisonic signal of order 3 has 16 input arguments plus the rotation angles and 16 output channels. While it is of interest to expose the Ambisonics order as an argument for the flexibility and reusability of code on the side of sclang, it is an argument unlikely to be changed while
an instance of the Ugen is running as a node in the DSP tree on scsynth. This is why for every order there is a unique Ugen for each function (encoding, transforming, decoding).

The Ugens follow these conventions, some of which are explained in subsequent sections):

- The Ambisonics channels are ordered according to the ACN convention.
- The default normalisation of the B-format is N3D.
- All azimuth and elevation arguments follow the spherical coordinates convention from SC.
- Operations of resource intensive Ugens can be bypassed.

Based on the implementation in Faust, the main functionalities of the HOA Ugens provided as the SC3plugins are so far:

- Encoding and decoding of planar waves and spherical waves using near field filters.
- Mirroring, Rotation (around azimuth and full 3D).
- Various Ugens for beamforming, returning mono as well as B-format signals.
- Various decoders in conjunction with Head-Related Impulse Responses (HRIRs) for binaural monitoring.

1.4.2 PseudoUgens

The second part of the library is available as the selang extension HOA Quark. While the SC3plugins are designed for computational efficiency of the sound synthesis processes, the HOA Quark is conceived to unlock the flexibility of making sound in SC with respect to code reusability and the scaling of synthesis scripts. Each typical operation in Ambisonics (Encoding, Transforming, Decoding) is here provided as a PseudoUgen. Depending on the Ambisonics order provided as an argument, the PseudoUgen returns and instantiates the correct Ugen from the SC3plugins collection on the sound server. Since the Ambisonics order is an argument for the PseudoUgen, the number of channels in the B-format vary and so does the number of input arguments in the Ugens. This is why the B-format is handled as a channel array. This makes the SC code flexible for experiments with different orders depending on computational resources. All arguments of the PseudoUgen obey the channel expansion paradigm. This means that if any of the arguments is an array (or an array of arrays), the PseudoUgen returns an array (or an array of arrays) of Ugens.

Figure 1 shows the relation between the SC language side (PseudoUgens in light grey) and the SC3plugins (Ugens dark grey). If the Ambisonics order is set to 3 and passed as an argument to the PseudoUgen, the corresponding Ugen with 16 channels is returned and a typical processing chain (Encoding Transforming Decoding, encircled in red) can be established. The main features of the design of the HOA library implementation on the language side are:

- B-format is handled as a channel array.
- All arguments obey the channel expansion paradigm.

This leads to the following advantage, when scripting HOA sound scenes in SC. Compared with graphical data flow programs like for instance Pure Data, changing the order means to reconnect all channels between objects interfacing with the B-format. In SC changing the Ambisonics order in a single global variable changes the order of the whole HOA processing chain.

1.4.3 Support directory

The third part of the library is a platform independent support directory for various HOA sound file recordings and convolutions kernels from HRIRs for the binaural rendering of HOA.
sound scenes. The support directory approach is similar to the Atk implementation from which we adapted the corresponding class. The reason to keep these resources separate from the Quark directory is mostly due to the size of the included sound files. We provide some 4th order HOA sound files that have been recorded with the Eigenmike® with the support of Romain Dumoulin from CIRMMT. The HRIRs provided are either measured from a KU-100 dummy head [Bernschuetz, 2013] or computed from a 3D mesh scan of several people’s face. The directions of the HRIRs follow a 50-node Lebedev grid, allowing an Ambisonic binaural rendering up to order 5 [Lecomte et al., 2016b].

2 Encoding

As the first step in an Ambisonics rendering chain, the library provides PseudoUgens for encoding into the B-format. One for the encoding of mono sound signals, one for microphone array prototypes and one for the commercially available Eigenmike® microphone array.

2.1 HOAEncoder

This PseudoUgen creates an HOA scene from mono inputs encoded as a (possibly moving) sound source in space. The source can be encoded 1) as a plane wave with azimuth and elevation ($\theta_p, \delta_p$) respectively 2) as a spherical wave with position ($r_s, \theta_s, \delta_s$), where $r_s$ is the distance to origin of the source. The spherical wave is encoded using near-field filters [Daniel, 2003]. In the current implementation, those filters are stabilized with near-field compensation filters. Thus, in this case, the radius of the loudspeaker layout $r_{spk}$ used for decoding is needed. Note that if the spherical source radius is such as the source is focused inside the loudspeaker enclosure ($r_s \leq r_{spk}$), a "bass-boost" effect may occur with potential excessive loudspeaker gain. This effect increases as the source get closer to the origin [Daniel, 2003] [Lecomte and Gauthier, 2015].

```
HOAEncoder.ar(1, SinOsc.ar(f),a,e); // returns [OutputProxy,...,OutputProxy]

HOAEncoder.ar(1, SinOsc.ar([f1,f2]), a,e); // returns
[[OutputProxy,...,OutputProxy], [OutputProxy,...,OutputProxy]]
```

If an array of azimuth and elevation arguments, matching in size those of the source $\text{SinOsc.ar}([f1,f2])$, flexible and scalable code for multi source encoding can be created.

2.2 HOAEncLebedev06 / 26 / 50, HOAEncEigenmike

This collection of PseudoUgen offers at first the Discrete Spherical Fourier Transform (DSFT) for various spherical layout of rigid spherical microphone. In the current implementation the proposed geometries are 06-26- or 50-node Lebedev grid [Lecomte et al., 2016b] and Eigen-Mike grid [Elko et al., 2009]. The components of the DSFT are then filtered to take into account the diffraction by the rigid sphere and retrieve the Ambisonic components [Moreau et al., 2006] [Lecomte et al., 2015] The filters are applied by setting the filter flag to 1 as shown in the next code listing:

```
// Encode the signals from the
// Lebedev26 grid microphone
HOAEncLebedev26.loadRadialFilters(s);
{HOAEncLebedev26.ar(4, SoundIn.ar(0!26), filters: 1)}.play
```

3 Converting

In order to correctly reconstruct a sound field from the channels of the B-format, it is important to know about standard normalization methods for the spherical harmonic components, as well as channel ordering conventions. Two main channel ordering conventions exist: The original Furse-Malham (FuMa) [Malham, 1999] higher-order format, an extension of traditional first order B-format up to third order (16 channels). FuMa channel ordering comes with maxN normalization, which guarantees maximum amplitude of 1. The FuMa format has been widely used and is still in use but is increasingly replaced by the Ambisonic Channel Number (ACN) ordering [Nachbar et al., 2011]. ACN typically comes with (the full three-D normalisation) where all signals are orthonormal. SN3D (Semi-Normalized 3D) spherical harmonics. This normalization has the advantage that
none of the higher order signals exceeds the level of the first Ambisonic channel, W (ACN 0).

However, this normalization does not provide an orthonormal basis of spherical harmonics and this latter case is recommended for transformations which rely on the orthonormality property of spherical harmonics. Therefore, the library uses internally the N3D (full 3D normalization) with ACN convention.

### 3.1 HOAConvert

The HOAConvert PseudoUgen accepts a B-format array as input and converts from and to ACN_N3D, ACN_SN3D, FuMa_MaxN. It is mostly meant to convert existing B-format recordings into ACN N3D for use within the library. The other use case is to render B-format mixes to other conventions for other production contexts.

### 4 Transforming

In its current implementation, the HOA library provides 3 standard operations like rotation and mirroring to transform the B-format.

#### 4.1 HOAAzimuthRotator

This PseudoUgen rotates the HOA scene around the z-axis, which is accomplished with a rotation matrix in x and y due to the symmetry in z of the spherical harmonics. For the matrix definition see [Kronlachner, 2014b]. In combination with horizontal head tracking, this transformation can stabilise horizontal auditory cues for left-right movements when the rendering is made over headphones in VR contexts.

#### 4.2 HOAMirror

This PseudoUgen mirrors an HOA scene at the origin in the directions along the axes left-right (y), front-back (x), up-down (z). According to [Kronlachner, 2014b], this can be accomplished by changing the sign of selected spherical harmonics.

#### 4.3 HOARotatorXYZ

This PseudoUgen rotates a HOA scene around any given angle around x,y,z. The rotation matrix is computed in spherical harmonic domains using recurrence formulas [Ivanic and Ruedenberg, 1996].

### 5 Beamforming

#### 5.1 HOAHCard2Mono

This PseudoUgen extracts a mono signal from the HOA scene according to a beampattern. The channels from the B-format inputs are combined to produce a monophonic output as if a directional microphone was used to listen into a specific direction in the sound field. In the current implementation, the beampattern provided are regular hypercardioids up to order 5 see [Meyer and Elko, 2002].

#### 5.2 HOAHCard2HOA

This PseudoUgen applies a hyper-cardioid beam-pattern to the HOA scene to enhance some directions and outputs a directional filtered HOA scene [Lecomte et al., 2016a]. The proposed beampatterns are regular hypercardioids as described in [Meyer and Elko, 2002]. The selectivity of the directional filtering increases with the order of the beam-pattern. This transformation requires an order re-expansion such that the output HOA scene should be of the order of the input HOA scene plus the beampattern order [Lecomte et al., 2016a].

#### 5.3 HOADirac2HOA

As in the previous section, this PseudoUgen performs a directional filtering on the HOA scene but this time the beam-pattern is a directional Dirac, that is to say a function which is zero everywhere except in the chosen direction. As a result the output HOA scene contains only the sound from the chosen direction. Thus, this tool helps to explore the HOA scene with a "laser beam". For more details see [Lecomte et al., 2016a].

### 6 Decoding

For the decoding of HOA signals two different ways of rendering the sound field are possible: First via headphones, or second through a setup of multiple loudspeakers.

For the headphone option the HOA signal is decoded to spherically distributed virtual speakers. For the best possible spatial resolution more speakers are needed than there are channels in the B-format. Each speaker signal is then convolved with HRTFs and the resulting left and right channels are summed respectively. For the distribution of the virtual speakers a regular distribution on the sphere is desirable, so that the decoding matrix is well behaved. This is why according to [Lecomte et al., 2015] and similar to the microphone array prototypes from above a Lebedev grid is chosen.
6.1 HOADecLebedev

This collection of PseudoUgen decodes an Ambisonics signal up to 50 virtual speakers positioned as nodes on a Lebedev grid. The decoding on the 50-node Lebedev grid works up to order 5. This grid contains two several nested sub-grids which work up to lower order with less nodes [Lecomte et al., 2016b]. Therefore, the first 6 nodes are sufficient for first order and the first 26 nodes are sufficient up to the third order. If the HRTF filter flag is set to 1, the signals are convolved with the kernels and summed up to yield a left and right headphone speaker signal. Prior to this, the convolution kernels need to be loaded to the sound server as shown in the next code listing:

```plaintext
// load a HOA sound file
~file = Buffer.read(s,"hoa30.wav");
// prepare binaural filters
HOADecLebedev26.loadHrirFilters()
{ HOADecLebedev26.ar(3, // order
    PlayBuf.ar(16,~file,1,loop:1),
    hrir_Filters:1)
}
```

6.2 HOADec

For the case of decoding for speaker arrays [Heller et al., 2008] distinguish 3 cases:

1. regular polygons (square, octagon) and polyhedra (cube, octahedron)
2. semiregular arrays (non equidistant but opposing speakers, like in a shoebox)
3. general irregular arrays (e.g. ITU 5.1, 7.1 ... semispherical speaker domes)

For the cases 1 and 2, decoder matrices can be obtained by matrix inversion. If, depending on the positions of the speakers, the resulting decoder matrix has elements of similar magnitudes, it is suitable for signal processing. For case 3, which are arguably the more realistic cases, a variety of state of the art techniques exists, see for instance [Zotter et al., 2012], [Zotter et al., 2010], and [Zotter and Frank, 2012]. An implementation of these techniques exceeds the scope of this library. However, for the construction of decoders for specific irregular speaker arrays, we refer the user to the Ambisonic Decoder Toolbox by Aaron J. Heller [7]. This toolbox produces decoders as Faust files, which can be compiled online as Ugens and in turn can then be integrated in the HOADec PseudoUgen class template.

7 The distance of sound sources

One novel aspect of the underlying Faust implementation of Ambitools is the spherical encoding of sound sources using near field filters. For the correct reproduction of the HOA scene, the distance of the sound source and the radius of the reproducing (virtual) speaker array needs to be set. The correct near field filters are either applied by setting it in the encoding or in the decoding step.

```plaintext
// load the binaural filters
HOADecLebedev26.loadHrirFilters()
{ var src;
    src = HOAEncoder.ar(
        3, // order
        PinkNoise.ar(0.1), // source
        az, // azimuth
        ele, // elevation
        plane_spherical:1, // set the speaker radius here
        radius:2,
        hrir_Filters:1)
    HOADecLebedev26.ar(3, // order
        src, // source
        // or set the speaker radius here
        // speaker_radius:1,
        hrir_Filters:1)
}
```

8 HOA and SynthDefs

The use of PseudoUgens leads to one important caveat when working with SynthDefs. The Ambisonics order is an argument pertaining to the PseudoUgen, it can hence not be an argument of a SynthDef. The reason is that at compile time the Ambisonics order would remain undefined and the PseudoUgen does hence not know which Ugen to return. When working with SynthDefs code reusability can still be achieved as shown in the next code listing:

```plaintext
// set the max order:
~order = 5;
~order.do({|i| // iterate
    SynthDef( // create unique names
        "hoaSin"++(i+1),asString,
```

https://bitbucket.org/ambidecodertoolbox/adt.git

http://faust.grame.fr/onlinecompiler/
9 HOA and Node Proxies

For the flexible creation of typical Ambisonics render chains, Node Proxies [Rohrhuber and de- Campo, 2011] provide an excellent tool in SC. Node Proxies autonomously handle audio busses and conveniently allow to crossfade between audio processes of a selected node, freeing silent process when the crossfade is completed. This allows to dynamically change sources in the encoding, transforming and decoding step in the rendering chain. The following code example shows a flexible scenario with changing seamlessly from an XYZ rotation to beamforming.

```plaintext
~o=3;
~chn=(~order+1).pow(2);
// load hoa sound file:
~bf=Buffer.read(s,"file.wav");
// b-format file player:
~player=NodeProxy(s,~audio,~chn);
~player.source=
{PlayBuf.ar(~chn,~bf)};
// Node for xyz rotation:
~trans=NodeProxy(s,~audio,~chn);
~trans.source=
{var in;in=in.ar(0!16)
HOATransRotateXYZ.ar(~o,in,
yaw,pitch,roll)};
// rotate the scene
~trans.set(~yaw, angle);
// decoding,
~dec=NodeProxy(s,~audio,~chn);
~dec.source=
{var in;in=in.ar(0!16)
HOADec.ar(~o,in,)};
// chain the proxies together
~player <>> ~trans <>> ~dec;
```

// change rotation to beamforming
~trans.source=
{var in;in=in.ar(0!16)
HOABeamDirac2Hoa.ar(~o,in,
az,ele)};
// direct the beam
~trans.set(~az, angle);

10 Conclusions

We have presented a HOA library for SC. The design of which resulted in great flexibility and makes it a valuable addition to experiment with HOA in various contexts. Due to the meta approach through Faust, future additions to the library are feasible and we look forward to experiment with it in the context of VR and video gaming platforms but also for the creation of sound material for electro acoustic compositions. We believe that the flexibility and live coding capacity of SC is particularly useful in the context of HOA, where repeated listening is essential to asses the perceptually complex mutual interdependence of temporal and spatial sound characteristics.

11 Acknowledgements

This work has been supported through the research-creation funding program of the Fonds de Recherche du Québec - Société et Culture (FRQSC).

References


M Kronlachner. 2014b. Spatial transformations for the alteration of ambisonic recordings. *Graz University Of Technology, Austria*.


