

Kurs: **DA1005 Självständigt arbete,  
komposition, 30 hp**

2014

Konstnärlig masterexamen i musik, 120 hp

Institutionen för komposition, dirigering  
och musikteori

---

Handledare: Henrik Frisk

David Granström

# Immersion

Spatialisering av elektronisk musik

Skriftlig reflektion inom självständigt, konstnärligt arbete

Det självständiga, konstnärliga arbetet finns  
dokumenterat på inspelning: Prism - David Granström

# Innehåll

<b>1</b>	<b>Introduktion</b>	<b>3</b>
<b>2</b>	<b>Spatialisering</b>	<b>3</b>
2.1	Immersion . . . . .	3
<b>3</b>	<b>Praktiska erfarenheter</b>	<b>4</b>
3.1	HTO . . . . .	4
3.2	György-Ligeti-Saal . . . . .	4
3.2.1	Undulation . . . . .	5
<b>4</b>	<b>Ambisonics</b>	<b>6</b>
4.1	Vad är ambisonics? . . . . .	6
4.2	Ordningar . . . . .	7
4.3	Varför använda sig av ambisonics? . . . . .	7
4.4	Att arbeta med Ambisonics . . . . .	8
4.5	Implementationer . . . . .	9
4.5.1	The Ambisonic Toolkit . . . . .	10
4.6	Transformationer . . . . .	11
4.7	Arbetsflöde . . . . .	11
4.7.1	DAW:s . . . . .	12
<b>5</b>	<b>Audiorama</b>	<b>12</b>
5.1	Beskrivning . . . . .	12
5.2	Spatialisering . . . . .	13
5.3	Examenskonserten . . . . .	13
<b>6</b>	<b>Aurora II</b>	<b>13</b>
6.1	Encoding . . . . .	13
6.2	Form . . . . .	14
6.3	Skiften mellan fixerat och generativt . . . . .	14

<b>7 Dispersion</b>	<b>15</b>
7.1 Form . . . . .	15
7.2 Konverteringar . . . . .	15
7.3 Decoding . . . . .	16
<b>8 Prism</b>	<b>16</b>
8.1 Form . . . . .	16
8.2 Tonhöjder . . . . .	17
8.2.1 Low . . . . .	18
8.2.2 Mid . . . . .	18
8.2.3 High . . . . .	18
8.3 Spatialisering . . . . .	19
<b>9 Diskussion</b>	<b>19</b>
<b>10 Referenser</b>	<b>20</b>
<b>11 Bilaga</b>	<b>21</b>

# 1 Introduktion

Den här texten är skriven som en summering över vad jag hittills har hunnit med att utforska vad det gäller virtuell/konkret spatialisering under mina år på KMH. Den är inte tänkt att fungera som någon generell översikt för spatialiseringstekniker i sig, även om dessa kommer att omnämnas till mer eller mindre hög grad genom texten. Den kommer heller inte att behandla spatialisering inom instrumentalmusiken. Jag kommer att ägna mest utrymme åt Ambisonics eftersom det är den teknik jag främst valt att arbeta med i min egen musik. För mig hör spatialisering, och då främst den immersiva effekten av spatialisering, nära ihop med den musik jag skriver. Jag upplever att statisk musik, t ex *drone*, i sig självt kan etablera ett klingande rum eller plats i det fysiska rummet (lyssningsplatsen). Spatialiseringstekniker kan förstärka upplevelsen ytterligare, och det är också därför jag har valt att utforska just detta fält i mitt konstnärliga arbete.

Texten är strukturerad i tre delar. Först en kort introduktion om spatialisering följt av praktiska erfarenheter, därefter kommer jag skriva något mer tekniskt om Ambisonics<sup>1</sup>. Slutligen kommer jag att skriva om min examenskonsert och om de tre verk som spelades där: *Aurora II*, *Dispersion* och *Prism*.

## 2 Spatialisering

När vi talar om spatialisering i elektroakustisk musik, eller annan sorts musik för den delen, så talar vi oftast om ljudens placering i det fysiska rummet. Det kan röra sig om musikers placering på en scen, i en konsertsal, eller olika typer av högtalaruppställningar och hur ljudmaterial är distribuerat i dessa. Förutom den del av spatialisering som är direkt knutet till rummet går det även att tala om *virtuell spatialisering*, som sker på en nivå innan ljudet når ut i det fysiska rummet. För min egen del är tankarna kring virtuell spatialisering det som intresserat mig mest vad gäller mitt eget komponerande. Jag ser oftast på mitt ljudmaterial som ett sorts rum, där ljuden samexisterar och etablerar en klingande plats. På den virtuella nivån av spatialisering har ambisonicstekniken varit till stor hjälp för mig, mycket tack vare det faktum att den per definition är fränkopplad den annars vanliga paradigmen en-signal-till-en-högtalare som återfinns i andra diskreta surroundformat som t ex 5.1 eller 7.1.

### 2.1 Immersion

Den aspekt som tilltalar mig mest med det som brukar kallas för surroundljud, eller helt enkelt för surround, är det som på engelska brukar benämnas som *immersion*, ett ord som ofta används i sammanhang för att beskriva den omslutande upplevelsen av virtual reality.

---

<sup>1</sup>I min kandidatuppsats *Auditiv rumslighet* behandlar jag de mer grundläggande aspekterna av Ambisonics.

Det svenska ordet immersion har egentligen samma betydelse – *att fördjupa sig i, att sänka sig ned i, att gå helt in i* – men det är kanske inte lika väl etablerat för att beskriva en sådan upplevelse som den engelska motsvarigheten gör i skrivande stund. När jag lyssnar på musik så blundar jag oftast eftersom jag vill att lyssnandet ska få ta så stor del av min uppmärksamhet som möjligt. Därför tycker jag även om att släcka ned belysningen på högtalarkonserter, eftersom det inte finns något att se. Jag upplever att surround kan vara till stor hjälp för att nå en högre nivå av immersion, eftersom det finns mer hörbar information att bearbeta när ljudet kan komma från alla riktningar i rummet. Upplevelsen av att befinna sig på en annan akustiskt plats än den faktiska lyssningsplatsen kan även förstärkas med hjälp av mer realistiska reverb som t ex ambisoniskt convolution reverb<sup>2</sup> som kan dra nytta av högtalare både under och ovanför lyssnaren.

## 3 Praktiska erfarenheter

### 3.1 HTO

På institutionen har vi haft tillgång till HTO, Högtalar Orkestern, som består av ett ständigt växande antal högtalare. I grunden har den bestått av 12 st Genelec 1030 högtalare samt 2 st Genelec subbasar. Den har nyligen vuxit med ett tillskott av 8 st L’Acoustic. Jag har varit med och arrangerat konserter med HTO och använt dessa högtalare på olika sätt under min utbildning. Förutom att ha arbetat med etablerade uppställningar som Quad, 5.1, 7.2, “Double Diamond” (8-kanaler) osv., har vi också arbetat mycket med okonventionella uppställningar. Något som jag upptäckte då och som jag sedan fortsatt att arbeta med är rumsreflektioner; att ställa upp och rikta högtalare mot t ex väggar eller andra objekt i rummet. Just för högtalardiffusion av stereoinspelningar i multikanalssetups lämpar det sig extra bra, eftersom det ger en så pass stor klangförändring av materialet utan att använda sig av t ex EQ eller andra effekter från mixerbordet. Resultatet kan kännas mer övertygande än om processeringen hade skett före output, just för att det är själva rummet som får stå för filtreringen av materialet vilket gör att det smälter samman på ett annat sätt med utljudet från de övriga högtalarna.

### 3.2 György-Ligeti-Saal

Jag reste till Graz, Österrike, februari 2013 för att vara med på IMPULS Academy – ett symposium för ny musik riktat till tonsättare. Jag hade fått en plats på en kurs i spatialisering (en del i större projekt kallat *Choreography of Sound*) med Gerhard Eckel som handledare. György-Ligeti-Saal är en nybyggd konsertsal som hör till KUG (Die Kunstuniversität Graz) i en byggnad som går under namnet MUMUTH. Konsertsalen är utrustad med bl a 33 st högtalare som är monterade på motordrivna stativ, med möjlighet

---

<sup>2</sup>Att konvolvera signalen med en impulsrespons i B-format. Se sektion 4 för mer information om B-format.

att kontrollera deras positioner vertikalt samt deras rotation i 360° i alla riktningar. Det enda som alltså inte går att justera är högtalarnas placering i förhållande till varandra. Deras placering i rummet utgår från en slumpad distribution som är speglad i rummets mittpunkt för att skapa en symmetri. Det går sedan enkelt att spara högtalarnas positioner och återskapa dessa från kontrollbordet i salen.

Arbetet var planerat så att vi först fick fundera på vad vi ville göra i rummet, vi fick sedan ett schema med tider då vi kunde testa och experimentera med våra idéer. Förutom jag själv var det även ca. 6 st andra kursdeltagare från olika delar av världen, alla tonsättare med en bakgrund i elektroakustisk musik. Min första tanke var att arbeta med Ambisonics och göra en perifonisk avkodning<sup>3</sup> med en hemisfärisk högtalaruppställning. Men eftersom jag redan visste att det skulle fungera (en sådan uppställning hade redan provats och kalibrerats i salen), ville jag ta chansen att utforska andra spatialiseringsmöjligheter. Jag valde därför att arbeta med högtalarna diskret – att behandla dem som enskilda ljudkällor i rummet i motsatts till ett ambisoniskt ljudfält – och istället använda mig av ambisonictekniker på en virtuell materialnivå. Rummet lämpar sig mycket bra för högtalardiffusering, eftersom att väggarna är konstruerade av ett relativt reflekterande material, samt att det finns två stora dörrar på en av kortsidorna som gav en annan klang än det övriga rummet om en högtalare riktades mot dem.

### 3.2.1 Undulation

Jag hade fått en slot på 2 x 30 minuter på en sorts öppen avslutningskonsert för den kurs jag hade deltagit i som jag fick disponera fritt. Jag bestämde mig för att skriva ett stycke som fick ta upp hela den tiden och fick arbeta intensivt för att hinna klart musiken i tid. Under kursveckan som var ca. sju dagar hade vi tillgång till salen vissa schemalagda tider, mycket av arbetet med musiken fick därför ske utanför konsertsalen. Som tur fanns det en mjukvara – *Virtual Mumuth* – utvecklad av IEM (Institut für Elektronische Musik und Akustik), vilken var till mycket stor hjälp. *Virtual Mumuth* använder sig av impulsresponser inspelade i salen på olika lyssningspositioner med en sorts “default” högtalaruppställning. Det går även att förflytta lyssningspositionen och det visualiseras även med en grafisk 3-dimensionell representation av rummet. Att sitta i den faktiska salen och arbeta med ljudmaterial i *Virtual Mumuth* med hörlurar på, var en nästintill skrämmande övertygande upplevelse.

*Undulation* är uppbyggt kring en långsamt glissando mellan två intervall,  $\frac{77}{73}$  – en mycket liten sekund och  $\frac{9}{8}$  – en ren stor sekund. Envelopen som styrde glissandot hade formen av en sinuskurva, den började på ratiot  $\frac{77}{73}$  och steg sedan till  $\frac{9}{8}$  efter hälften av styckets totala duration (15 minuter i det här fallet), och sedan ned igen under styckets andra hälft. I den långa versionen av stycket som är 60 minuter, byts  $\frac{9}{8}$  istället ut mot  $\frac{321}{256}$  – i närheten av en tempererad stor ters. Det fanns en grundton på 174.61 Hz (ett F) som var utspridd i

---

<sup>3</sup>3-dimensionell avkodning. Avkodningar kommer att förklaras mer i detalj i sektion 4.

hela rummet och klingade konstant genom hela stycket som en drone. Jag hade placerat högtalarna i rummet som en hemisfär där det rörliga intervallet förflyttade sig genom olika högtalare i en bana i formen av en dubbelhelix genom rummet. När det (intervallet) återvänt till sin ursprungliga position i rummet triggades ett perkussivt ljudobjekt med mycket energi i det lägre frekvensregistret som en sorts påminnelse över att en cykel i rummet var gjord. Antalet cykler var kontrollerbart och räknades ut baserat på den **totala durationen / antal cykler**. I versionen jag spelade upp vid konserttillfället var stycket segmenterat i 14 cykler på 2 minuter och 8 sekunder vardera. Förutom detta material fanns även ett annat basmaterial som triggades halvvägs in i en cykel, en lägre sinusaktig bas som spelade grundtonen en oktav ned. Det fanns även ännu ett perkussivt material i form av filtrerade impulser, som slumpades någon gång under varje cykel och som spred ut sig i rummet på olika sätt. Impulsmaterialet kunde börja längst bak i salen och röra sig framåt igenom den, eller det motsatta, det kunde även förflytta sig i ett sorts sicksacksmönster mellan högtalarna, samt slumpas på olika platser i rummet.

## 4 Ambisonics

### 4.1 Vad är ambisonics?

Ambisonics är ett samlingsnamn över en serie upp- och inspelnings tekniker som möjliggör två-dimensionell (pantofonisk) och tre-dimensionell (perifonisk) ljudåtergivning genom ett obestämt antal högtalare. Det som särskiljer ambisonics från andra surroundformat är att all spatial information finns inkodad i en signal som kallas för *B-format*. B-formatsignalen kan sedan avkodas till konkreta högtalaruppställningar, eller till hörlurar (binaural). Det är viktigt att förstå att vi inte längre talar om enskilda högtalare när vi avkodar en B-formatssignal. Man anger för avkodaren hur många högtalare som finns till förfogande, och eventuellt högtalarnas positioner i rummet genom att specificera *azimuth* (horisontell riktning) och *elevation* (vertikal riktning), detta görs vanligtvis i måttenheten radianer. Tillsammans hjälps sedan högtalarna åt för att återge det som kallas för ett ambisoniskt *ljudfält*. Ett ljudfält går att visualisera som en sfär där ljudmaterial kan placeras ut på punkter eller områden på sfären (ljudfältet), det går även att tänka sig att ljudmaterialet “belyser” olika områden av sfären beroende på ljudets placering. Det går att tänka på högtalarna som en sorts motor som genererar ljudfältet och ju fler det finns, desto mindre behöver enskilda högtalare arbeta för att återge den avkodade signalen. Vissa menar att fler billigare högtalare kan låta “bättre” om de används för ambisonisk avkodning än vad färre men dyrare högtalare skulle gjort, då det totala antalet högtalare gör att de enskilt inte behöver arbeta lika hårt för ambisonisk ljudåtergivning.<sup>4</sup>

De teoretiska och praktiska aspekterna av Ambisonics utvecklades på av bl a Michael Gerzon vid Oxford Mathematics Institute under 1970-talet i England. Tekniken slog aldrig

---

<sup>4</sup>Samtal med Joseph Anderson under en workshop om Ambisonics och *ATK* i Karlsruhe 2013.

igenom kommersiellt, och har för det mesta använts inom akademiska sammanhang och av personer med specialintressen för inspelningstekniker. Ambisonics har dock fått större uppmärksamhet under de senaste åren, bland annat för att t ex datorspelsindustrin och virtual reality utvecklare har börjat intressera sig allt mer för tekniken.

## 4.2 Ordningar

Ambisonics är ett hierarkiskt format där B-formatsignalen kvalificeras i olika ordningar. Det går att tänka på ordningar i ambisonics som en sorts spatial samplingsfrekvens, där högre ordningar ger en mer detaljerad spatial representation eller upplösning. Vi skiljer oftast på det som kallas *first order ambisonics* (FOA) och alla ordningar över den första brukar gå under samlingsnamnet *higher order ambisonics* (HOA). I första ordningen representeras B-format av fyra diskreta kanaler som kallas WXYZ. Den första ordningen är egentligen en sammansättning av ordning 0 och 1, och på samma sätt ingår ordning 0, 1, 2 och 3 i den tredje ordningen osv. Summan av antalet kanaler i B-formatet för en viss ordning kan uttryckas som  $N = (M + 1)^2$  (perifonisk återgivning), eller  $N = 2M + 1$  (pantofonisk återgivning) där  $N$  är antalet kanaler i B-formatet och  $M$  är ambisonicsystemets ordning (Hollerweger, 2008). Att antalet kanaler i B-formatssignalen ökar betyder också att antalet högtalare som behövs för att korrekt kunna reproducera ljudfältet ökar. Det teoretiska antalet högtalare som behövs är  $L \geq N$  där  $L$  är antal högtalare och  $N$  är antalet kanaler i B-formatsignalen. Om antalet högtalare inte är tillräckligt för att avkoda ett format från den högre ordningen går det bra att utelämna kanalerna från den högre ordningen och istället avkoda de som får plats i den givna högtalaruppställningen. Då systemet är hierarkiskt är högre ordningar alltid bakåtkompatibla med lägre, där den lägsta nivån (ordning noll) då skulle vara monofonisk.

## 4.3 Varför använda sig av ambisonics?

Bland det första som tilltalade mig med att använda Ambisonics som spatialiseringsteknik var separationen av den spatiala informationen kontra den faktiska högtalaruppställningen. Jag hade tidigare upplevt en begränsning över att veta att den musik jag skrev för t ex Quad- eller 5.1 uppställningar var så att säga "låst" vid dessa format. Självklart kunde jag mixa ner de diskreta kanalerna till t ex en stereorepresentation, men det var något jag ofta drog mig från att göra eftersom jag sällan blev särskilt nöjd med resultatet. Ambisonics blev en lösning på detta dilemma, eftersom jag säkert visste att all den spatiala informationen fanns inkodad i B-formatmixen och att avkodaren skulle göra sitt bästa för att återge det ljudfält jag hade skapat oberoende av antal högtalare.

Enligt mig är den absolut största fördelen med ambisonics **portabilitet**. Det faktum att det enkelt går att spela upp en B-formatmix i en mängd olika typer av högtalaruppställningar utan att behöva göra något särskilt med mixen i sig känns oerhört tilltalande. Att inte



behöva tänka på högtalare, och istället helt fokusera på de spatiala idéerna i kompositionen har gett mig mycket mer frihet i mitt tänkande kring spatialisering.

Det leder mig till den andra riktigt stora fördelen med ambisonics gentemot diskreta surroundformat, nämligen att det går att manipulera ljudfältet med spatiala filtreringar. Att göra något, konceptuellt enkelt, som att flytta alla ljudkällor i en mix och rotera dem i någon riktning, är otroligt enkelt i jämförelse med diskreta system. Det räcker att manipulera själva *ljudfältet*, det finns ingen anledning att manipulera enskilda signaler i mixen för att försöka få dem att rotera i samma hastighet och riktning. Exempel på klassiska ambisoniska transformationer är *rotate*, *tilt* och *tumble*. Där 'rotate' applicerar en rotation av ljudfältet kring z-axeln, 'tilt' roterar ljudfältet kring x-axeln medan 'tumble' applicerar en rotation kring y-axeln. Dessa transformationer kan även med fördel kombineras.

En vanlig kritik riktad mot ambisonics är att punktkällor inte känns lika "direkta" som vid diskret panorering. Tänker man på det så är det heller inte så konstigt, det är svårt att slå den spatiala "cue" som en faktiskt konkret ljudkälla (t ex en högtalare) i ett rum utgör. Min erfarenhet av ambisonics har främst varit i den första ordningens B-format, men jag har tidigare haft tillfälle att lyssna på upp till tredje ordningen, och redan där märks en stor skillnad i riktningsverkan vid punktljudkällor. I min egen musik lider jag inte särskilt av detta faktum, då jag mest intresserar mig för den immersiva effekt som ett ambisoniskt ljudfält kan skapa.

Något annat som kan kritiserar är vikten av en symmetrisk högtalaruppställning. Detta är vanligtvis inte ett problem, men det kan ändå få konsekvenser om det t ex inte skulle gå att justera de högtalare som finns på plats i en konsertlokal. Min erfarenhet är dock, och det kan mycket väl bero på mitt musikaliska material, att även om det blir något sämre spatial återgivning inte blir fullständig katastrof om högtalarna inte står på exakt rätt positioner i förhållande till varandra. Om det finns tid och utrustning så går det även att kalibrera utgångssignalen med hjälp av delay (om högtalare behöver komma "längre bort" från centrum) och tillika amplitudkompensation.

#### 4.4 Att arbeta med Ambisonics

Att börja arbeta med Ambisonics kan till en början kännas främmande, kanske till stor del för att tekniken frångår de mest vanliga surroundtekniker och deras paradigmer. Men i sin allra enklaste form går det utan problem att använda sig av en helt vanlig monosignal som kodas om till B-format, och som i sin tur kan avkodas till högtalare. Att vara medveten om de olika "stegen" som signalen kan befinna sig när man arbetar med Ambisonics är viktigt för att förstå vart i signalkedjan man bör göra vad. Förutom att koda om monosignaler, som kan vara elektroniskt framställda eller inspelade med en mikrofon, går det också att använda sig av specialiserad utrustning som t ex *Soundfieldmikrofonen* eller *TetraMic*. Soundfieldmikrofonen är uppbyggd av fyra stycken subkardiod, eller kardiod kapslar, som sitter mycket nära varandra (i teorin skulle de helst befinna sig på exakt samma punkt)

i formen av tetraeder. Signalen från de fyra mikrofonkapslarna går under benämningen *A-format*, vilken inte är tänkt att användas i sin råa form. Istället kodas A-formatet om till B-format, som därefter går att använda för fortsatta ambisoniska processeringar. På detta sätt är det alltså även möjligt att framställa syntetiskt A-format genom att använda sig av fyra stycken dekorrelerade signaler av samma ljudkälla. Det enklaste exemplet skulle vara fyra signaler av t ex dekorrelerat brus som sedan kodas om till B-format med en lämplig matris. Jag använder nästan uteslutande den här tekniken när jag arbetar med syntes i Ambisonics, eftersom det ger en mycket stor flexibilitet vad gäller transformationer, samt att det ger upphov till en större immersiv effekt efter avkodning. Det liknar fenomenet, eller känslan av djup, som uppstår av att ha två dekorrelerade signaler utplacerade i vardera högtalare i ett stereofält, mot att panorera en monosignal till mitten av fältet vilket kan upplevas som plattare.

Formeln här ned visar matrisen för det första steget som behöver tas för att konvertera mellan A- till B-format. A-formatet utgår från fyra stycken mikrofonkapslar (fysiska eller virtuella) ordnade i formen av en tetraeder (som Soundfieldmikrofonen) där LFU, RFD, RBU och LBD står för “left-front-up”, “right-front-down” etc. Men för att få en korrekt B-formatsignal (WXYZ i första ordningen), behöver resultatet från matrisen även filtreras och skalas om (Fons, 2007).

$$\begin{aligned}
 W' &= LFU + RFD + RBU + LBD \\
 X' &= LFU + RFD - RBU - LBD \\
 Y' &= LFU - RFD - RBU + LBD \\
 Z' &= LFU - RFD + RBU - LBD
 \end{aligned}$$

## 4.5 Implementationer

De flesta patent som rör Ambisonics har numera utgått, även det som rör Soundfieldmikrofonen, vilket gör tekniken möjlig att implementera för vem som helst. Många av implementationerna som finns idag är fri mjukvara, vilket innebär att källkod finns tillgänglig samt att programmen oftast är gratis att använda. Jag har dels arbetat med VST/AU plugins som kan användas i DAW:s, men främst har jag arbetat med ett bibliotek till programmeringsmiljön SuperCollider som heter “The Ambisonic Toolkit”. Bland de VST/AU plugins jag använt känns det värt att nämna “Ambisonic Studio B2X plugins” utvecklade av Daniel Courville, som är samling encoders/decoders samt effekter från första- och upp till femte ordningens Ambisonics. Andra populära plugins som också kan nämnas är bl a “WigWare”, “AmbiX” samt “Harpex”.

### 4.5.1 The Ambisonic Toolkit

The Ambisonic Toolkit (ATK) har utvecklats av Joseph Anderson sedan 1998 och har under åren varit tillgängligt i olika former under ett antal plattformar som t ex Csound och Common Lisp Music. I sin nuvarande form är ATK implementerat som ett bibliotek för SuperCollider. Tanken bakom ATK är att ge användaren möjlighet att interagera med ambisonictekniken på ett sätt som ger full kontroll över alla steg i produktionen av ljudfältet. ATK uppmanar till att “tänka ambisoniskt”, på allt som sträcker sig bortom den mest vanliga paradigmen av surround – placering av ljudkällor i ett rum.

ATK har antagit följande modell som arbetsflöde för Ambisonics:

Author → Image → Monitor

- Author

*Att med hjälp av mikrofoninspelningar eller syntes skapa ett ambisoniskt ljudfält.*

ATK erbjuder många intressanta alternativ för encoding till B-format:

- Planewave: Classic directional encoding
- Omnidirectional: A soundfield from everywhere
- Virtual loudspeaker array: Transcoding standard formats
- Pseudoinverse microphone array: Encoding from discrete microphones or signals

- Image

*Spatiala filtreringar av ett ambisoniskt ljudfält.*

ATK tillhandahåller de mest klassiska spatiala Ambisonics transformationerna samt en del mer experimentella. De flesta av transformationerna finns tillgängliga både som statiska och dynamiska:

- Rotation: Soundfield rotation about an axis
- Mirror: Soundfield reflection across an axis
- Directivity: Soundfield directivity
- Dominance: Adjust directional gain of soundfield
- Focus: Focus on a region of a soundfield
- Push: Push a soundfield in a direction

- Monitor

*Att spela upp eller rendera ett ambisoniskt ljudfält.*

Några exempel på decoders som finns tillgängliga i ATK:

- Stereo UHJ: Classic Ambisonic stereo decoding
- Binaural: Measured and synthetic HRTFs

- Regular 2D & 3D: Single and dual polygons
- Diametric 2D & 3D: Flexible semi-regular arrays
- 5.0: Wiggins optimised decoders

Genom att ha denna paradigm som en sorts karta över arbetsflödet blir det ganska enkelt att komma igång och arbeta med Ambisonics. Det som gör att ATK sticker ut i jämförelse med många andra ambisonicsimplementationer som jag hittills stött på är att biblioteket är oerhört väldokumenterat med många exempel, samt att stor vikt läggs vid spatiala transformationer.

## 4.6 Transformationer

Att arbeta med transformationer, eller spatiala filtreringar som de också kan kallas, är något unikt för ambisonicstekniken. Jag har särskilt tyckt om att jobba med dem i syntesalgoritmer, där de kan kombineras och användas på sätt som de kanske inte från början var tänkta att användas till. Ett enkelt exempel skulle kunna vara att amplitudmodulera signalen med hjälp av t ex en rotations-transformation. Mer involverade exempel kan t ex vara att applicera statistiska transformationsmatriser i en rundgångsloop. Att arbeta med transformationer på detta och liknande sätt går även att se som ett exempel på virtuell spatialisering, där arbetet med riktningar inte nödvändigtvis behöver “höras” i det konkreta rummet.

Jag tänkte ge ett exempel på en intressant ambisonisk processering som jag använt mig av och fick lära mig av Joseph Anderson under en workshop om ATK i Karlsruhe. Processeringen går ut på att först applicera en kontinuerlig rotation av ljudfältet och sedan avkoda det till A-format, processera signalen i A-format (med t ex filter, distortion, reverb etc.), och sedan konvertera signalen tillbaka till B-format och applicera motsvarande rotation i samma hastighet men då i motsatt riktning. Resultatet blir att *processeringen* förflyttar sig runt i ljudfältet, men själva ljudfältet verkar stanna på samma plats. Den här typen av transformation är något som gör ambisonicstekniken mycket användbar för mitt eget konstnärliga skapande. Jag arbetar ofta med statistiska klanger i min musik, och detta blir ett praktiskt sätt att introducera variation i ett annars statistiskt ljudmaterial men samtidigt kunna behålla den övergripande statistiska kvaliteten.

## 4.7 Arbetsflöde

Att arbeta med Ambisonics i post-produktionsfaser som mastering kräver en något annorlunda metod än den man annars skulle tillämpa i mastering av diskreta surroundformat. Det som är viktigt att komma ihåg är att alla former av dynamisk processering, så som kompression, distortion etc, *inte* går att applicera direkt på B-formatssignalen utan att den spatiala informationen skadas. Linjära och tidsinvarianta processeringar som equalization

går däremot bra att applicera direkt på B-formatssignalen, förutsatt att den påverkar alla kanaler på samma vis. Ett exempel skulle vara ett notch-filter på 1000Hz som appliceras över alla WXYZ kanaler för en första ordningens B-formatssignal. Men det går att arbeta med dynamiska effekter trots detta, lösningen är att avkoda B-formatet till A-format (som är likvärdig signalen från mikrofonkapslarna i en Soundfieldmikrofon), efter att processeringen har skett på A-formatssignalen kan vi återigen koda om den till B-format med vår spatiala information intakt. När jag har arbetat med post-produktion av min egen musik har jag avkodat B-format till A-format i non-realtime i SuperCollider och sedan arbetat med de resulterande ljudfilerna i Reaper (DAW).

#### 4.7.1 DAW:s

Att arbeta med Ambisonics kräver vissa förutsättningar vad det gäller routing hos en DAW (Digital Audio Workstation). Av de mjukvaror jag provat har jag kommit fram till att *Reaper* i dagsläget har de bästa routingmöjligheterna för att producera musik i Ambisonics. Jag har även hört bra om *Ardour*, en fri mjukvara som först och främst utvecklats för Linux, men som även finns tillgängligt för OS X. De stora aktörerna Avid's *Pro Tools* samt Apple's *Logic Pro*, har begränsade routingmöjligheter i jämförelse och fokuserar snarare på kommersiella surroundformat som 5.1 eller 7.1.

Genom att använda Reapers grupp- och routingsystem skapade jag först ett 4-kanals spår för A-formatljudfilen. Sedan skickade jag utsignalen från kanalerna 1-2 (WX) och kanalerna 3-4 (YZ) till två separata stereo spår. Dessa spår var sedan grupperade under ytterligare ett 4-kanals spår som reglerade volymen för spår WX samt YZ och som slutligen skickades till masterbussen. Genom att arbeta på detta sätt kunde jag enkelt lyssna och konfigurera vanliga stereoeffekter genom att först lägga till en A-B encoder på WXYZ kanalen och sedan efter det en UHJ (stereo) decoder. När jag hade ställt in mina plugins som jag ville ha dem, var det enkelt att bara kopiera över dessa på WX- respektive YZ-kanalerna och ta bort eller bypassa alla encoders/decoders på WXYZ gruppen. Efter att ha exporterat ljudfilerna från Reaper konverterade jag tillbaka dem till B-format genom samma non-realtime process i SuperCollider. Det kan låta omständligt, men efter ett att hittat en arbetsstruktur som fungerade var det inte särskilt tidskrävande <sup>5</sup>.

## 5 Audiorama

### 5.1 Beskrivning

*Audiorama* är en scen för akusmatisk musik, hörspel och ljudkonst. Den kupolformade lokalen är belägen på Skeppsholmen i Stockholm. Audiorama är utrustat med 21 högtalare som är utplacerade i hela rummet. Det finns en cirkel av åtta högtalare, fem stycken inuti

<sup>5</sup>Se bilaga för en grafisk representation av signalfödet.

valvet i kupolen ovan, fyra stycken vid golvhöjd, samt fyra stycken subbasar utplacerade i vardera hörn av rummet. Rummet är mycket bra akustiskt kalibrerat, med otroligt tydlig definition och i princip ingen efterklang. Lite som att spela i en välbyggd studio, men med 55 publikplatser i mitten av rummet.

## 5.2 Spatialisering

Med beskrivningen här ovan kan man tänka sig många olika typer av spatialiseringsmöjligheter. Allt i från att komponera i 21 stycken diskreta spår, till att utnyttja alla högtalare som ett ambisoniskt ljudfält. Jag har under mina år på kandidatprogrammet, och nu även masterprogrammet, fått chansen att arbeta med rummet på Audiorama på olika sätt. Ett av mina tidigare flerkanalstycken *Dispersion* skrevs specifikt för Audiorama.

## 5.3 Examenskonserten

Jag hade min examenskonsert på Audiorama den 11:e april 2014, där även mitt examensstycke *Prism* uruppfördes. Jag valde att använda mig av en perifonisk avkodningsmatris för alla tre stycken som spelades under konserten, vilket innebar att alla 21 högtalare i rummet utnyttjades.

# 6 Aurora II

*Aurora* är ett generativt stycke som jag skrev hösten 2013. Jag gjorde en ny version inför konserten, där jag även skrev om syntesalgoritmen till Ambisonics som innan var skriven i stereo. Den nya tagningen fick heta *Aurora II*. De olika versionerna skiljer sig ganska mycket från varandra eftersom många parametrar i musiken baseras på slump – även vissa formgivande delar.

## 6.1 Encoding

Jag valde att skriva om syntesalgoritmen till A-format. A-formatet går sedan enkelt att koda om till B-format med hjälp av en A-B matris. Det var inte helt enkelt att göra detta, eftersom att syntesen till stor del är beroende av rundgång, samt att alla olika syntinstanser är ihopkopplade med varandra spektralt. Det tog alltså ett ganska stort mått av experimenterande för att nå fram till ett godtagbart resultat som kunde mäta sig med den första versionen av stycket. Ett annat problem som dök upp var att syntesen blev så pass mycket tyngre CPU-mässigt för datorn att rendera. Detta är kanske främst ett problem som kan lösas med hjälp av snabbare hårdvara eller optimeringar i syntesalgoritmen. SuperColliders syntesserver arbetar dessutom bara på en tråd, så det går tyvärr inte att tillgodogöra sig kraften från flerkärniga processorer. En alternativ syntesserver med stöd

för multi-threading är under utveckling – *supernova* – vilket kommer förbättra situationen avsevärt (Blechmann, 2010). Men det är ändå något att vara medveten om vid refaktorering av syntesalgoritmer från t ex stereo till Ambisonics. Detta var bara första ordningens Ambisonics, så det rör sig bara om fyra kanaler (dvs två extra från stereo). Men att skriva om en algoritm för t ex tredje ordningen skulle alltså innebära totalt 16 kanaler för ett perifoniskt ljudfält. Det är blir en väldigt stor skillnad i antal uträkningar.

## 6.2 Form

Jag skrev *Aurora* mycket snabbt, under en kortare period där jag försökte skriva ungefär ett nytt stycke i veckan. Formen är således inte särskilt överarbetad, men det var också själva tanken med kompositionen, att överlägga mycket av de formbärande materialet till slump. Det jag själv har bestämt vad gäller formen, är att det kommer ett skifte efter ca 1/3 av stycket längd, där kontrollsignalen som styr spektrala förändringar i syntesen separeras så att de olika lagren istället kontrolleras individuellt. Det som faktiskt spelar roll i formupplevelsen, åtminstone i min mening, är vilka bastoner som väljs och vilken intensitet dessa får. Detta är alltså inget jag valt att kontrollera mer än att jag bestämt vilka toner det faktiskt får slumpas mellan. På grund av detta, har jag känt det nödvändigt att spela in olika tagningar av stycket för att jämföra dessa sida vid sida, och sedan välja ut “den bästa”. De är alla ganska snarlika, som urvalsprocess bestämde jag mig för att om jag tyckte om början och slutet betydde det att det var en okej tagning.

## 6.3 Skiften mellan fixerat och generativt

Det sker ett skifte i mitt lyssnande när jag har spelat in ett generativt stycke musik till en fixerad inspelning. Plötsligt är musiken en ljudfil som går att hoppa runt i, med tydligt start och slut. Jag tänker nästan aldrig på musiken så förrän jag väl sitter där med den nyinspelade musiken i en mixningssession. Plötsligt blir detaljer som jag inte brytt mig så värst mycket om viktiga, samma sak med styckets duration. Jag skulle inte säga att jag lyssnar mindre aktivt när jag arbetar generativt, snarare är det så att jag lyssnar på ett annat sätt. Jag flyttar fokus till områden som känns viktigare för min lyssningsupplevelse när jag vet att musiken inte har någon början eller något slut och potentiellt skulle kunna pågå i en evighet.

När jag hamnar i dessa situationer, där jag plötsligt behöver kontroll över en detaljnivå som inte finns tillgänglig direkt i min källkod, eller som kanske är mycket svår att uppnå utan att skriva om stora delar av programmet, brukar jag tillämpa två metoder: (A) Försöka begränsa utfallet av eventuell slump som används och genom iterativ design komma fram till slumpvärden som “alltid låter bra”. Eller (B) separera olika lager och spela in dem individuellt för att sedan arbeta med dem för hand i en DAW. Alternativ ‘B’ kan ibland vara mer problematiskt, eftersom det kan uppstå en konflikt i förhållningen till det strikt

algoritmiskt genererade och mina egna intuitiva idéer om materialet. Därför har jag valt att använda mig av detta alternativ under vissa begränsningar. Jag flyttar inte om ljud på tidsaxeln utan låter dem vara kvar på de platser där de först genererades, och jag klipper heller inte bort något i materialet. Genom att ha den här förhållningen till post-produktion av generativ musik, behöver jag inte ställas inför alla de val som potentiellt skulle gå att göra. Jag har istället fattat mina val före jag påbörjar arbetet. Det jag däremot brukar göra är volymjusteringar på macro- och micronivå i de olika lagren samt processeringar med plugins där jag tycker att det behövs.

## 7 Dispersion

*Dispersion* skrevs 2011 inför en konsert på Audiorama i samband med KDM institutionens festival *ljudOljud*. Jag hade aldrig varit på Audiorama innan och därför fick jag arbeta med spatialiseringen i musiken rent konceptuellt. Jag hade en idé om att byta plats i höjddled på två olika skikt i musiken, en långsam transition som knappt skulle vara märkbar om man bara lyssnade efter den, men som kompositoriskt hörde starkt ihop med styckets form.

### 7.1 Form

Jag använde mig av en övergripande intensitetskurva när jag arbetade med formen. Det blev enkelt att representera i SuperCollider som en envelope. När jag bestämt hur kurvaturen och hållpunkterna i envelopen skulle se ut, duplicerade jag den i tre individuella instanser, som alla hade en liten slumpad variation mot varandra i maximal intensitet och tidsangivelser för envelopehållpunkterna. Jag mappade sedan dessa tre enveloper på olika parametrar i mitt material och skapade på så sätt en gemensam rörelse men med individuell variation mellan lagren. Styckets totala duration var skalbar och jag kunde därför lyssna och justera längden under tiden jag komponerade.

### 7.2 Konverteringar

Ett problem jag direkt ställdes inför efter konserten, var hur jag skulle göra en stereomix av de 17 diskreta kanaler jag hade arbetat med på Audiorama. Det kom fram en person till mig som undrade om jag kunde skicka en stereoversion av stycket, som en demo, eftersom hen arbetade med en filmskapare som eventuellt kunde vara intresserad av min musik. Veckan som följde gjorde jag en downmix där jag helt enkelt panorerade ut materialet i stereofältet baserat på de ursprungliga positionerna av högtalarna på Audiorama, samt volymjusterade och filterade de olika spåren. Jag skapade senare ytterligare en version i Quad. Quadversionen är den jag använt vid senare uppspelningstillfällen.



## 7.3 Decoding

När jag spelade *Dispersion* på examenskonserten använde jag mig inte av originalversionen för Audiorama utan istället Quadversionen. Det var dels ett beslut baserat på tidsbrist, eftersom jag inte hade tillgång till de ljudfiler jag exporterat för 17 kanalsversionen, men det var även på grund av att jag ville prova att transkoda Quadformatet till Ambisonics. Jag kunde lyssna på skillnaderna mellan de olika versionerna mycket enkelt i Studio A på KMH där jag arbetade med åtta högtalare, liknande den åtta-ring som finns på Audiorama. Resultatet blev lyckat, men samtidigt något av en kompromiss. Jag uppskattade Quadversionen något bättre om jag satt på *sweet spot* mitt i rummet och lyssnade. Men för en publik så är det kanske på sin höjd 3-4 personer som skulle kunna få den lyssningsupplevelsen som jag hade i studion. Därför beslöt jag mig för att använda mig av den Ambisoniska versionen istället, eftersom jag där hade möjlighet att spatialt filtrera ljudfältet med hjälp av *near field compensation*. Det är en spatial filtreringsteknik som virtuellt förflyttar högtalarna längre bort från mittpunkten i rummet, och efter att ha suttit på olika kanter i rummet och lyssnat bestämde jag mig för att det skulle bli den bästa återgivningen för flest antal personer på konserten.

## 8 Prism

När jag inledde arbetet med *Prism* var det till en början långt ifrån självklart vad för sorts musik jag hade tänkt skriva, och det tog mig ganska lång tid tills jag hade förstått vilka parameterar i stycket som var viktigast. Jag har insett att jag är känslig för, och bunden till det klingade materialet. Utan den typen av feedback blir det oftast långa omvägar och tvära stopp i mitt arbete. I arbetet med *Prism* var jag tvungen att ändra mitt arbetsflöde på grund av att jag inte kunde generera det klingade materialet i realtid, och var istället tvungen att rendera det i non-realtime för att sedan kunna lyssna på det. Dessa extra steg gjorde kompositionsprocessen något längre än vad jag först hade väntat mig, och ändringar i materialet tog längre tid att utvärdera eftersom det involverade flera led av export.

### 8.1 Form

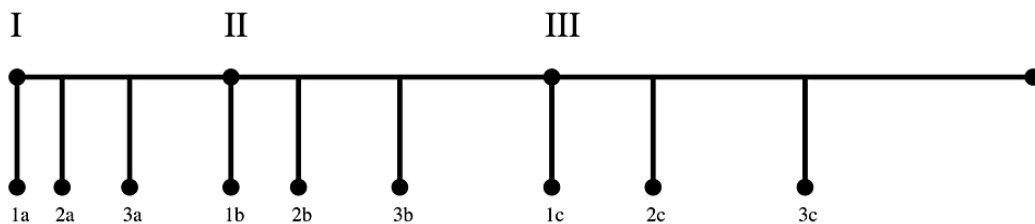
Jag hade formen klart för mig nästan från början. Jag ville använda intervall för att beskriva durationer. I *Prism* är alla durationer konstruerade kring ratiot  $\frac{3}{2}$ , 1.5 eller med andra ord intervallet en ren kvint. Formen är uppbyggd på så sätt att jag tagit ratiot  $\frac{3}{2}$  upphöjt med sig självt tre gånger, och dessa tre delar utgör formen för huvudsektionerna i stycket. Jag normaliserade formdelarnas värden så att summan av dem blev 1, och skalade sedan dessa värden med styckets totala duration. På detta sätt kunde jag experimentera mer fritt med den totala durationen utan att låsa fast mig vid en viss total duration innan allt material fanns på plats.

```

// Define a function 'f' which takes three arguments
f = {arg numParts, ratio, totalTime;
  var x;
  // Collect the ratios in an array
  x = numParts.collect {|i| pow(ratio, (i+1)) }; // [ 1.5, 2.25, 3.375 ]
  // Normalize the array so that its sum equals 1
  x = x * x.sum.reciprocal;
  // Multiply the normalized array with the number of seconds
  x = x * totalTime;
  // Return the result
  x;
};
f.value(3, 3/2, 60 * 17.5); // [ 221.05263157895, 331.57894736842, 497.36842105263 ]

```

Funktionen här ovan är skriven i SuperCollider och visar hur jag räknade ut längderna för de olika huvudsektionerna i stycket. Jag tog sedan samma funktion och applicerade den över varje huvudsektion. Resultatet blev då tre stycken subsektioner inom varje huvudsektion med samma proportioner som storformen.



Figur 1: *Prism - form*

## 8.2 Tonhöjder

Jag arbetar nästan uteslutande med det som brukar kallas ren intonation eller 'just intonation' i min musik. Att arbeta med ren intonation ger i princip obegränsade möjligheter till att stava om, och utforska intervall som inte finns tillgängliga i tempererad stämning. Jag valde ut ett antal ration som fick representera en tonal bank för det här stycket  $\frac{1}{1}$   $\frac{3}{2}$   $\frac{4}{3}$   $\frac{5}{3}$   $\frac{5}{4}$   $\frac{9}{8}$   $\frac{15}{8}$  sedan tillkom även oktaveringar av dess ration. Tonhöjdernas placering i stycket valdes olika beroende på de olika lagren. Lagren var uppdelade i tre stämmor, som jag kallade *Low*, *Mid*, *High* baserat på register. Här följer en översikt över hur det tonala materialet var distribuerat genom stycket:

### 8.2.1 Low

Basen spelade en fixerad melodi (cantus firmus). Varje subsektion av de tre delarna har antingen en tonhöjd eller en paus.

Del	Ratios
1	<i>Paus</i> $\frac{15}{8}$ $\frac{5}{4}$
2	$\frac{5}{3}$ <i>Paus</i> $\frac{5}{3}$
3	$\frac{3}{2}$ $\frac{15}{8}/2$ $\frac{1}{1}$

Tabell 1: Low

### 8.2.2 Mid

Mellanstämman spelade duoler enligt ett förutbestämt rytmiskt mönster. Samma mönster som den översta stämman fast i retrograd.

Del	Ratios
1	$[\frac{2}{1} \frac{3}{2}]$ $[\frac{1}{1} \frac{2}{1}]$
2	$[\frac{2}{1} \frac{15}{8}]$ $[\frac{2}{1} \frac{5}{4}]$ $[\frac{1}{1} \frac{5}{3}]$
3	$[\frac{3}{2} \frac{5}{4}]$ $[\frac{5}{3} \frac{15}{8}]$ $[\frac{2}{1} \frac{5}{3}]$

Tabell 2: Mid - Duoler

### 8.2.3 High

Den översta stämman spelade antingen duoler eller fyrklanger slumpat.

Del	Ratios
1	$\frac{5}{4}$ $\frac{9}{8}$ $\frac{5}{3}$ $\frac{15}{8}$
2	$\frac{1}{1}$ $\frac{2}{1}$ $\frac{3}{2}$ $\frac{5}{4}$ $\frac{5}{3}$ $\frac{15}{8}$
3	$\frac{1}{1}$ $\frac{5}{4}$ $\frac{3}{2}$ $\frac{2}{1}$

Tabell 3: High - Duol till fyrklang.

### 8.3 Spatialisering

Jag ville framför allt att fokus skulle hamna på klangen och formen i *Prism*. Därför valde jag att använda mig av rätt så begränsade spatiala filtreringar. Det lager som utmärker sig mest skulle nog vara det tredje (High), där varje enskild ton börjar på en punkt i rummet och som sedan “sprider ut sig” och blir jämt fördelad över ljudfältet. De andra lagren är statiska, förutom ett svagt brus som går genom ett convolution reverb, en bakgrund som skapar större rymd i vissa delar av stycket. Bruslagret var filtrerat genom en långsam rotate-tilt-tumble transformation<sup>6</sup>.

## 9 Diskussion

När vi skriver elektronisk musik ger valet av teknik alltid vissa möjligheter men även begränsningar – spatialiseringstekniker ej undantaget. Valet av spatialiseringsteknik kan ha implikationer på det konstnärliga resultatet, vilket kan vara medvetet eller omedvetet från tonsättarens sida (Baalman, 2010). Därför tycker jag att det känns viktigt att fråga sig varför man valt en viss uppsättning verktyg eller tekniker i sitt konstnärliga arbete. Kommer de att finnas kvar 10 år framåt i tiden? Om inte, vilka alternativ finns då tillgängliga? Framför allt tror jag det är viktigt att hitta en balans i att hinna lära sig bemästra en teknologi fullt ut, men samtidigt inte måla in sig i ett hörn om det skulle visa sig att verktygen skulle bli oanvändbara av kompatibilitetsskäl eller liknande i framtiden. För att lättare kunna fatta dessa beslut har jag valt att se på verktyg och teknologier ur ett sorts empiriskt perspektiv. Ungefär hur många använder sig av verktyget eller tekniken idag? Hur länge har verktyget eller tekniken funnits? Om det är en mjukvara, är det fri programvara med öppen källkod eller kommersiell? Hur lätt skulle jag själv kunna implementera tekniken i fråga om verktyget jag använder mig av idag skulle försvinna? Jag har inget emot att använda mig av nya och experimentella tekniker i mitt skapande – tvärt om. Men jag ser det som mycket viktigt att ha en stabil grund att utgå från, något som inom den elektroniska musiken inte är lika självklart som papperet och pennan skulle kunna vara för den notbaserade konstmusiken.

Spatialisering på någon nivå, virtuell eller reell, är något alla som arbetar med elektronisk musik kommer att stöta på och behöva hantera. Att skapa djup i en stereomix eller arbeta med flerkanalig musik för en stor uppsättning högtalare kan båda vara exempel på en problematik som eventuellt kan avhjälpas genom att göra sig medveten om att det handlar om en spatial problematik. Det finns många andra mer eller mindre moderna spatialiseringstekniker som inte är omnämnda i denna text som t ex *VBAP*, *DBAP* och *Wave Field Synthesis*. Jag skulle gärna ägna dessa teknologier större uppmärksamhet i framtiden, och på så vis utvidga min palett av tekniker för spatialisering.

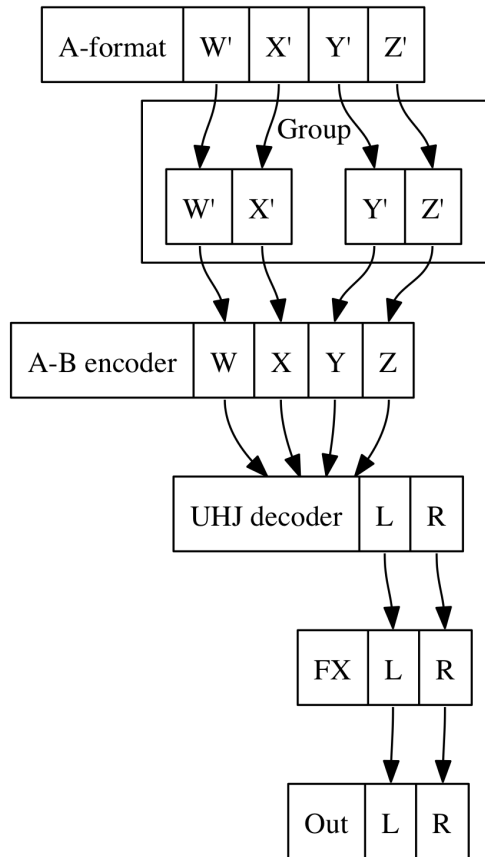
---

<sup>6</sup>Se sektion 4.3 för en närmare beskrivning av dessa transformationer.

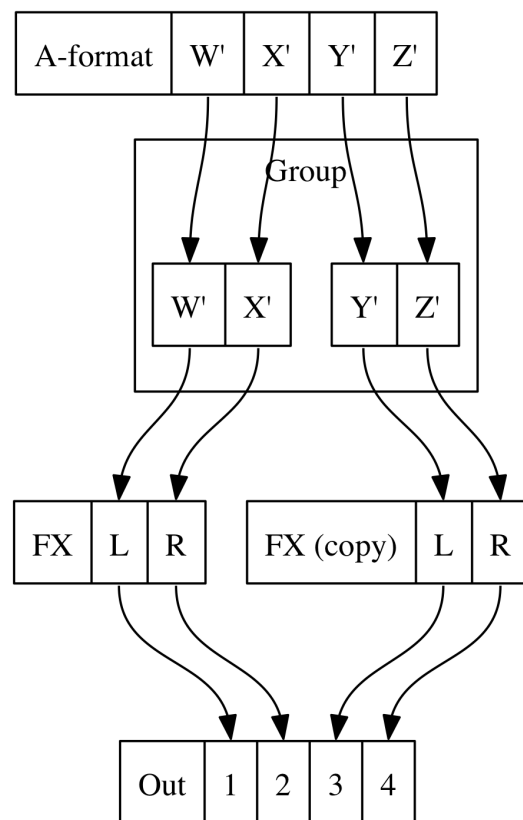
## 10 Referenser

- Adriaensen, F. (2007). A tetrahedral microphone processor for ambisonic recording. *Proceedings of the Linux Audio Conference 2007*. Retrieved from <http://kokkinizita.linuxaudio.org/papers/tetraproc.pdf>
- Baalman, M. (2010). Spatial composition techniques and sound spatialisation technologies. *Organised Sound 15(3) 2010*.
- Blechmann, T. (2010). Supernova, a multiprocessor-aware synthesis server for SuperCollider. *Linux Audio Conference 2010*. Retrieved from [http://tim.klingt.org/publications/lac2010\\_supernova.pdf](http://tim.klingt.org/publications/lac2010_supernova.pdf)
- Hollerweger, F. (2008). An introduction to higher-order ambisonic. Retrieved from <http://flo.mur.at/writings/HOA-intro.pdf>

# 11 Bilaga



Figur 2: *A-format FX pre-listen*



Figur 3: *A-format FX render*

