Veirt le Blaungk

A Complete Technical & Theoretical Analysis

Code Architecture · Infinitesimal Voice Leading · Scriabin's Mystic Tetrachords

Spiral Time Structures · Formant Synthesis · Bayle's Acousmonium

Lawrence McGuire

2025-2026

Contents

§1 Introduction

The 10 minutes of voice music finds its base in the rarely heard work by composer and post-spectralist Kaija Saariaho. Following a desire to hear this piece started after encountering an online journal article from The Society of Music Theory. Here I found the score for "Vers le Blanc" and some accompanying audio fragments found online ([?]). Her idea on infinitesimal voice leading was a concept that aligned closely to my expereiments with working with hybrids of voice and the auditory tone perception of vowels. Especially the latter is something that Saariaho clearly wanted to implement in her infinitesimal voice leading. Does a perception of a clear vowel—or even more fundamental—a voice happen only when certain formant spacings, individual intensities and spectral bandwidths are achieved, or is there a region in which this percept is clearly there. Saariaho works with various elements in her long interpolation, not only timbre. For instance, tone change and especially their combined harmonic movements between pitch sets are very much explored in this work. I decided to also work with specific pitch class sets that play with Scriabins' mystical hexachords[1] and a derived version for the four voiced piece I have here. Not having heard the full piece beforehand was a blessing, and my main concern was to combine some ideas I had on the role of onsets in voice perception, from silent to voiced. Instead of leaving it at the Saariaho like uninterrupted tri-phonic tone/formant interpolations, I decided to introduce clear cuts in each section. The piece consists of 10 sections of $60s$. Each section either cuts up the sustained voice tones in: spiral aggregates (cfr. Clarence Barlow's 'spiral melodies'), pulsed gated structures, or more monophonic overlapping segments. Or the section is completely sustained.
The opportunity came to present this piece on Francois Bayle's Acousmonium which then prompted me to also commit to also a spatial considerations and essentially its diffusion into a space, in this instance it was in Amare's Conservatoriumzaal in The Royal Conservatoire of The Hague.

§2 Vowel Morphings

2.1 Formant Dictionaries

The core acoustic model treats the human voice as a bank of *four bandpass resonators* (formants), each defined by a centre frequency f_k and a reciprocal bandwidth (rQ) q_k. The dictionaries encode empirically measured formant data for eight standard IPA vowels in both male and female registers:

$$\text{Vowel } V \;\mapsto\; \big\{(f_1, q_1),\ (f_2, q_2),\ (f_3, q_3),\ (f_4, q_4)\big\} \tag{1}$$

For example, the male vowel `\a` has formant centres at 730 Hz, 1090 Hz, 2440 Hz, 3400 Hz with rQ values 0.11, 0.08, 0.05, 0.04 (narrower bandwidth = smaller rQ in SuperCollider's `BPF/Formant` UGen convention). The vowel target sequence used in the final system is:

[`rando`, `a`, `u`, `rando3`, `o`, `i`, `rando2`, `aw`]

Three stochastic entries (`rando`, `rando2`, `rando3`) generate random formant configurations at evaluation time via `exprand` and `rrand`, providing aperiodic spectral excursions that

[1] quartal hexachord consisting of an augmented fourth, diminished fourth, augmented fourth, and two perfect fourths.

punctuate the otherwise vowel-anchored sequence. Together, they create a timbral arc: open resonance → noise → closed front vowel → open back vowel. Interestingly, movements between random formant structures and vowels operate perceptually different. So, from `rando` to `a`, f_1 goes from 3134 Hz to 730 Hz and from q_1 from 0.0699 to 0.11. The amount of pitch change over 60s can be given by a delta/difference value $\lambda_{f_1} = 2404$ Hz and bandwidth change by $\lambda_{q_1} = -0.0401$. Then, f_3 and q_3 have respective lambdas λ_{f_3}, λ_{q_3} equal to 843 Hz and 0.239. The amount with which the third formant moves in respect to the rate of change with the first formant is much slower. This means that approximately the third formant will surpass a *critical distance* earlier at which it can be in, to still play its spectral constructive part in the overal vowel structure. And over the different formant interpolations, random-to-vowel or vowel-to-random movements will create complex because of different formant centers surpassing at different timings these critical distances.[2] There is sort of cascading effect when each formant starts to reach these critical distances. It is important however to understand the importance of $F1$, $F2$, $F3$ and $F4$. Some studies say that one only needs the first two formants to accurately perceive a vowel, others state the importance of the $F3$ but deem $F4$ to be negligible. There are different ways of "suddenly hearing a vowel" when coming from a random formant structure: if the $F1$ and $F2$ surpass their critical distances first, and if $F3$ and $F4$ arrive later, then their presence or contribution to the timbre perception is probably diminished or possibly disregarded as artefacts/side-products before eventually crossing the critical boundary. If $F3$ and $F4$ have arrived at their target way before the $F1$ and $F2$ have, then one can imagine how different the higher formant's roles are in the perceptual construction of the vowel.

2.2 Vowel Morphing via `IEnvGen`

Vowel morphing is realised through `IEnvGen` (indexed envelope generator) rather than pairwise `Select`/`Lag` interpolation. The entire vowel sequence is encoded as a multi-breakpoint envelope; a time-pointer (`Line.ar`) scans through it at audio rate:

```
var formantFreqs = freqArrays.collect { |levels|
    IEnvGen.ar(Env(levels, durs, \exp),
               Line.ar(0, totalDur, totalDur))
};
var bandwidths = rqArrays.collect { |levels, i|
    var rqEnv = IEnvGen.ar(Env(levels, durs, \lin),
                           Line.ar(0, totalDur, totalDur));
    formantFreqs[i] * rqEnv;
};
```

Listing 1: Formant frequency envelopes using `IEnvGen`

Formant frequencies use *exponential* curvature (`\exp`), reflecting the approximately logarithmic nature of psychoacoustic frequency perception (mel/Bark scales). Reciprocal-Q values use linear interpolation. Bandwidth is defined absolutely as $B_k = F_k \cdot \mathrm{rQ}_k$, so the filter aperture widens proportionally at higher formant frequencies — consistent with the physical behaviour of vocal-tract resonances.

[2] Formant frequency discrimination studies using well-trained listeners under optimal conditions establish that the second formant $F2$ can deviate approximately 1-2% from its center frequency while maintaining clear vowel perception.https://doi.org/10.1121/1.2029821

Interpolation equation. At any moment $t \in [0, T]$, the formant frequencies follow:

$$\hat{f}_k(t) = (1-\alpha)\, f_k^{(n)} + \alpha\, f_k^{(n+1)}, \qquad \alpha = \{t/T \cdot (N-1)\} \pmod 1 \tag{2}$$

where $\{\cdot\}$ denotes fractional part. The eight-state timbral trajectory weaves recognisable vowels (`a`, `u`, `o`, `i`, `aw`) together with three stochastic excursions (`rando`, `rando2`, `rando3`).

§3 Pitch Layer: Scriabin Mystic Tetrad Harmonic Interpolation

```
~mysticTetrad = { |root|
    root + [0, 6, 10, 16]
};

~clusters = [
    ~mysticTetrad.(48)/4,     // opening: very low register
    ~mysticTetrad.(42),
    ~mysticTetrad.(50),
    ~mysticTetrad.(45),
    ~mysticTetrad.(53),
    ~mysticTetrad.(47),
    ~mysticTetrad.(55),
    ~mysticTetrad.(40),
    ~mysticTetrad.(52),
    ~mysticTetrad.(44),
    ~mysticTetrad.(48)/10     // closing: descending into sub-bass
];
```

Listing 2: Mystic tetrad construction and cluster sequence

The pitch engine then interpolates between successive cluster states using `EnvGen` with a curve value of -4 (mildly convex):

$$p_v(t) = \text{EnvGen}\big(\text{Env}(\{c_v^{(0)}, \ldots, c_v^{(K)}\},\ \{T/K, \ldots\},\ -4)\big) \tag{3}$$

where $v \in \{0, 1, 2, 3\}$ indexes each of the four chord voices independently. The four resulting pitch envelopes become the *fundamental frequencies* of four independent `Formant.ar` units. T is $600s$ and K is the amount of sections equal to 11.

The result is a continuous, slowly-evolving harmonic landscape that passes through eleven mystic-tetrad states over the full 600 seconds, and directs us to a computational analogue of the Saariaho/Callender model of *infinitesimal voice leading* but then applied to Scriabin's harmonic language.

§4 The `Formant.ar` UGen: Coupling Pitch and Timbre

SuperCollider's `Formant.ar` is a band-limited impulse-train generator with built-in formant resonance. Its three parameters are:

- `fundfreq`: the fundamental F_0 (sets the spectral comb spacing)
- `formfreq`: the resonant peak frequency F_k

- `bwfreq`: the bandwidth of the resonant peak (in Hz)

For each of the four pitch voices v and each of the four formants k:

$$y_{v,k}(t) = \texttt{Formant.ar}\Big(F_0^{(v)}(t) + \text{tremolo}_k(t),\ F_k(t),\ B_k(t)\Big) \tag{4}$$

The final signal for voice v is the sum over all four formants:

$$y_v(t) = \sum_{k=0}^{3} y_{v,k}(t) \ \cdot\ G_k(t) \tag{5}$$

where $G_k(t)$ is the mute gate for formant channel k.

§5 Mute-Gate Architecture: The `~generateMutePatterns` Function

This is the main time structuring device. It generates four independent *amplitude gate envelopes*—one per formant channel—by dividing the total duration into `numSections` sections and assigning each section one of four behavioural modes chosen at random:

5.0.1 Mode A — \allOn

All formants are held continuously at full amplitude throughout the section. This creates moments of maximum spectral density, where all four resonance bands are simultaneously active.

5.0.2 Mode B — \overlap

Formant channels alternate in a staggered on/off pattern. Each channel is offset by a phase of $\phi_k = k/4$ of the segment, so channels enter and exit in succession. The inter-onset times are drawn from an exponential distribution (`exprand`), creating irregular, overlapping bursts of spectral energy:

$$G_k^{(\text{overlap})}(t) = \big\lfloor (n_t + \phi_k \cdot N) \bmod 2 \big\rfloor, \quad \phi_k = k/4 \tag{6}$$

5.0.3 Mode C — \pulse

High-rate gating (20–400 events per section) creates a rapid amplitude modulation. With short inter-onset intervals drawn from `exprand(0.3, 1.5)`, this produces effects ranging from slow tremolo to audio-rate amplitude modulation (AM) sidebands. Each formant channel starts on a different phase ($\phi_k = k$), so the four channels pulse in quadrature, producing a kind of *spectral tremolo* where different resonance bands are successively emphasised.

5.0.4 Mode D — \spiral

```
{ behavior == \spiral } {
    var numHits    = rrand(6, 14);
    var numRepeats = rrand(3, 8);
    var segmentLength = sectionDur / numRepeats;
    var baseIOIs = numHits.collect { exprand(0.3, 2.0) };
```

```
    baseIOIs = baseIOIs.normalizeSum * segmentLength;

    numRepeats.do { |repeatIdx|
        var evolutionType = [0, 1, 2].wchoose([0.7, 0.7, 0.7]);
        var evolution = linlin(repeatIdx, 0, numRepeats-1, 1.0,
            case
                { evolutionType == 0 } { rrand(0.5, 0.8) }   // shorten
                { evolutionType == 1 } { rrand(1.2, 1.6) }   // lengthen
                { evolutionType == 2 } { 1.0 }
        );
        var evolvedIOIs = (baseIOIs * evolution).normalizeSum
                            * segmentLength;
        sectionTimes = sectionTimes ++ evolvedIOIs;
    };
};
```

Listing 3: Spiral rhythmic evolution within mute patterns

In this piece, the random sequence was,

`[Spiral-Pulse-Sust-Overlap-Overlap-Sust-Sust-Pulse-Spiral-Sust]`

as can be seen in Figure 1. It takes a base rhythmic pattern (a sequence of inter-onset intervals, or IOIs) and *evolves* it through successive repetitions by scaling the IOIs by a factor that varies linearly from 1.0 to some target value (either shrinking $\in [0.5, 0.8]$ or expanding $\in [1.2, 1.6]$). Because each iteration is renormalised to the segment length, the overall rhythmic density stays constant even as the internal proportions shift:

$$\mathrm{IOI}_j^{(r)} = \frac{\delta^{(r)} \cdot \mathrm{IOI}_j^{(0)}}{\sum_j \delta^{(r)} \cdot \mathrm{IOI}_j^{(0)}} \cdot L_{\mathrm{seg}} \tag{7}$$

where $\delta^{(r)} = \mathrm{linlin}(r, 0, R-1, 1.0, \delta_{\mathrm{target}})$ is the per-repeat evolution scale factor.

This creates a *Fibonacci-like compression or expansion* of rhythmic patterns — the same structural idea as Saariaho's rhythmic interpolations in *Jardin secret II* (as described by Morrison), but operationalised as a self-contained procedural generator.

§6 Tremolo Architecture

Four independent tremolo oscillators, one per formant channel, are generated with slightly randomised frequencies around 6.25 Hz:

$$f_{\mathrm{trem},k}(t) = \mathrm{clip}(6.25 + 1.5 \cdot \xi_k(t),\ 1,\ 9) \qquad \xi_k \sim \texttt{LFNoise1.kr(0.05)} \tag{8}$$

Each tremolo is applied only when the corresponding mute gate is open (via an `EnvGen.asr` gated by a `Trig1` derived from gate transitions), and modulates the fundamental frequency of the corresponding `Formant.ar`:

$$F_0^{(v,k)}(t) = p_v(t) + 10\,\mathrm{Hz} \cdot \sin(2\pi f_{\mathrm{trem},k}\, t) \cdot G_k^{\mathrm{VCA}}(t) \tag{9}$$

The VCA envelope `tremoloVCA` is a step-function randomly toggling between 0 and 1 at 60-second intervals, adding a further layer of macro-level activation.

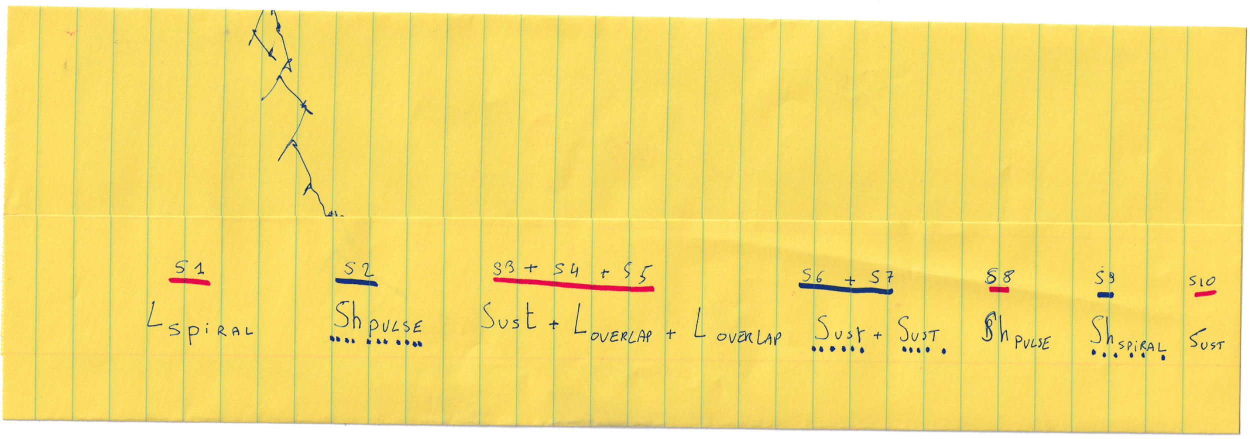

Figure 1: Modes per Section for "Veirt le Blaungk" (2026)

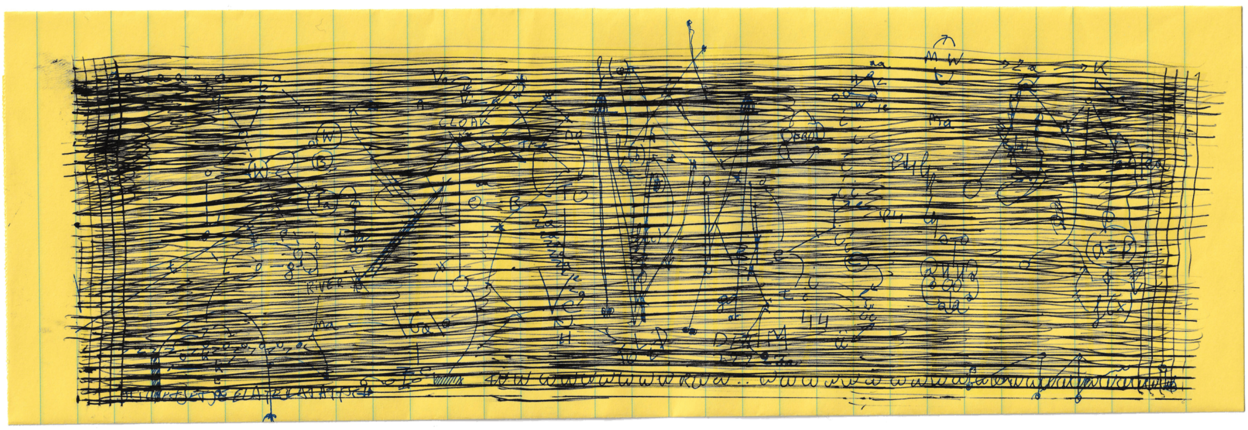

Figure 2: Visual Palate Cleanser

§7 Spatial Output

My initial plan for the Acousmonium was to send the four formant oscillators to certain *groups* of sort-of stereo speaker pairs of which there were plenty. The Acousmonium usually has a frontal focused speaker setup. The set-up in my mind before I received the actual speaker set-up in the concert hall can be seen in Figure 3, and the actual set up is in Figure 4.

Eventually the setup also included two rings of 8 speakers— one at ear level and one on the balcony—one quadraphonic setup which added some possibilities but also difficulties to my initial four groups idea. I decided to then make a stereo mixdown as I felt this four channel idea was already making things harder for myself. The Acousmonium had this time, minus the sub pair, 19 stereo pairs (one pair is not drawn here). So, a `Splay.ar` UGen mixes the four voices with slowly rotating stereo width (`LFPar.kr(1/150)`) and slowly drifting centre position (`LFPar.kr(1/125)`), with period ratios intentionally incommensurate (150 s vs. 125 s) to avoid periodic spatial repetition.

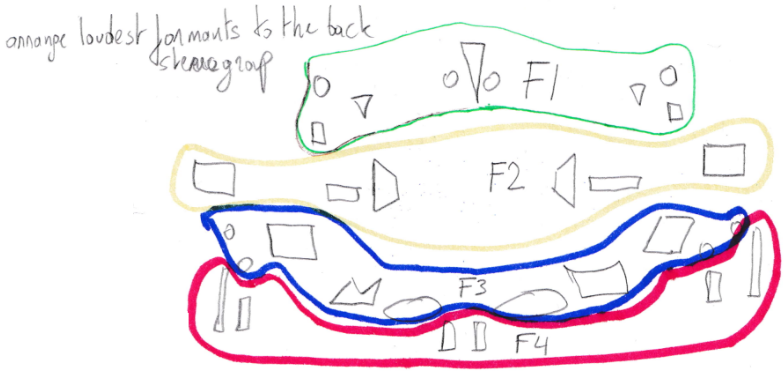

Figure 3: Groups of speaker pairings for each formant oscillator F^1,F^2,F^3 and F^4

§8 Infinitesimal Voice Leading

8.1 Callender's Framework

Clifton Callender's 2004 paper "Continuous Transformations" provides the mathematical foundation for understanding both the Saariaho original and this reimagining [?]. The central insight is that voice leading need not move in discrete semitone steps: it can proceed *infinitesimally*, tracing a continuous trajectory through a multi-dimensional pitch space.

Definition (Infinitesimal Voice Leading). Let $A = \langle a_1, \ldots, a_n \rangle$ and $B = \langle b_1, \ldots, b_n \rangle$ be ordered pitch sets. The continuous transformation

$$f(t) = \langle f_1(t), \ldots, f_n(t) \rangle, \qquad f_i(t) = (b_i - a_i)\, t + a_i \tag{10}$$

traces a straight-line path in n-dimensional Euclidean space from A to B. At no point along this path need $f(t)$ coincide with any equal-tempered pitch set.

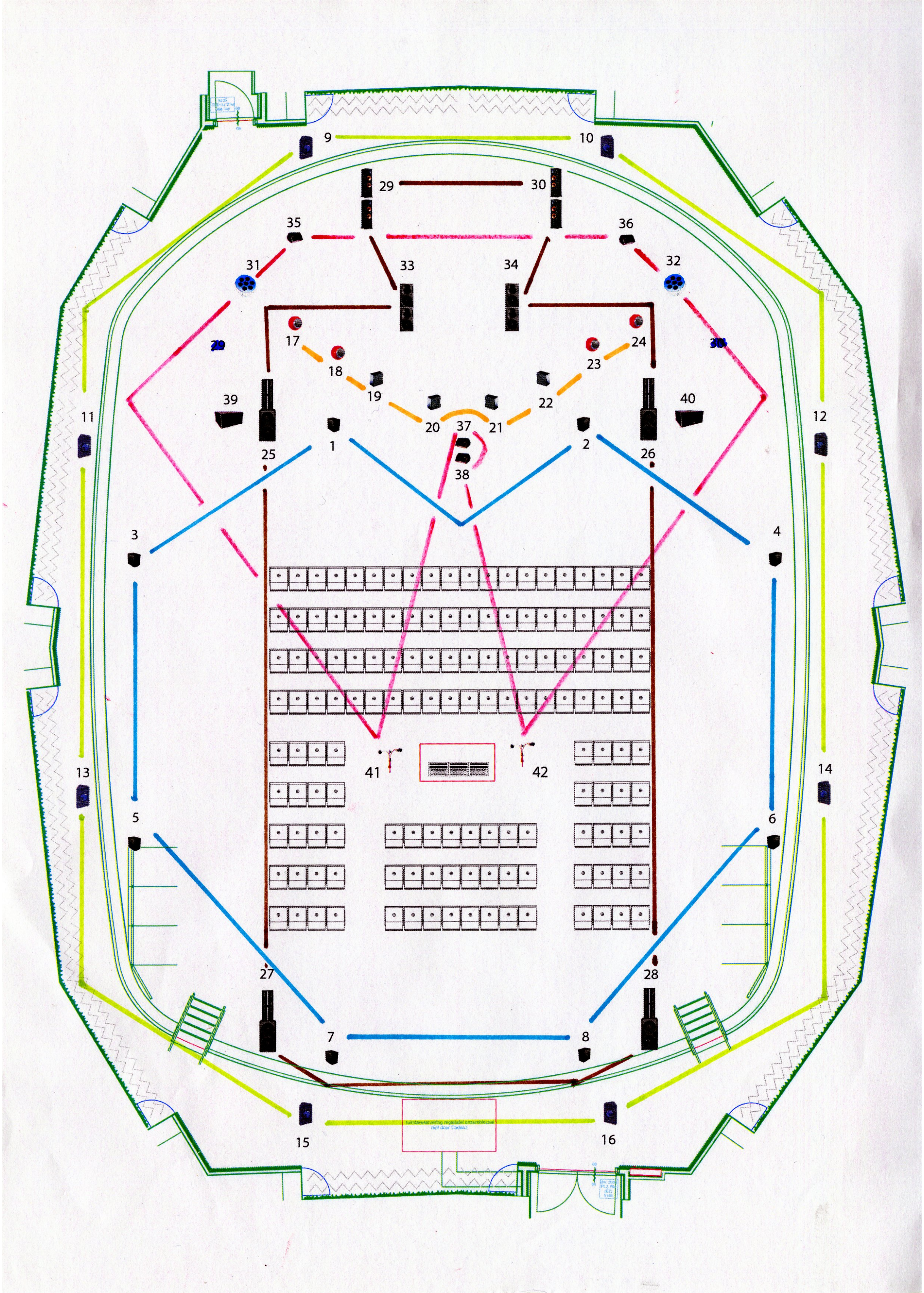

Figure 4: Actual Acousmonium set up in the concert hall

For *Vers le blanc* specifically, Callender identifies the three voices as:

$$f_1(t) = 4t \tag{11}$$
$$f_2(t) = -7t + 9 \tag{12}$$
$$f_3(t) = -6t + 11 \tag{13}$$

with $A = \langle 0, 9, 11 \rangle$ and $B = \langle 4, 2, 5 \rangle$ (C3 = 0).

8.2 T-class Space

To compare voice-leading paths modulo transposition, Callender constructs a *T-class space* using an oblique coordinate system. For three-voice sets, the class representative is $\langle a_2 - a_1,\ a_3 - a_1 \rangle_T$, and the distance between T-classes is:

$$\rho(A, B) = \sqrt{\frac{3}{2}\left(\sum_{i=1}^{3}(b_i - a_i)^2 - \frac{\left(\sum_{i=1}^{3}(b_i - a_i)\right)^2}{3}\right)} \tag{14}$$

For n voices, this generalises to:

$$\boxed{\rho(A, B) = \sqrt{\frac{n}{n-1}\left(\sum_{i=1}^{n}(b_i - a_i)^2 - \frac{\left(\sum_{i=1}^{n}(b_i - a_i)\right)^2}{n}\right)}} \tag{15}$$

This metric satisfies the key property: if A and B differ by the movement of a *single* voice by distance h, then $\rho(A, B) = h$.

8.3 Boundaries of the Fundamental Region

The T-class space is partitioned into a *fundamental region* Π^3 by three families of mirror boundaries corresponding to permutation, inversion, and modular equivalence. As the path $f(t)$ crosses a permutation boundary (e.g., $v_1 = v_2$), we hear a *voice crossing* — two pitch streams converge on a unison and then exchange relative positions. This is the most audible moment in *Vers le blanc* (occurring at $t = 9/11$, approximately 12'16" into the 15-minute work).

Key perceptual consequence. Reflections of $f(t)$ off the boundaries of Π^3 cause the trajectory to *retrace* nearby harmonic territory, producing passages of near-stasis where the set-class changes very little despite continuous motion. Callender quantifies this: between t_2 and t_4 (spanning a 36-second window centred on the permutation boundary), the distance traversed in Π^3 is only $\rho \approx 0.082$ — approximately a twelfth of a semitone of set-class displacement.

8.4 The Morrison Archival Perspective

Landon Morrison's 2021 article adds an essential dimension: the realisation that *Vers le blanc*, as notated and as theorised by Callender, conceals an entire secondary layer of *timbral* interpolation operating below the level of pitch. Morrison's access to the IRCAM archives revealed:

1. **Formant interpolations**: Each of the three synthesised voices passes through a continuous sequence of CHANT phoneme states (vowels and consonants), with breakpoint functions governing the temporal evolution of each formant region's amplitude and centre frequency.
2. **Rhythmic interpolations**: Amplitude peaks of the continuous voices are modulated by rhythmic patterns ($\{0.5, 0.33, 0.25\}$ = eighth, triplet eighth, sixteenth notes) that interpolate toward a triple subdivision by the end.
3. **Palindromic macro-structure**: The eight phonemic processes (I–VIII) are arranged in a loose mirror symmetry around the midpoint of the piece, creating a metaprocess that subsumes all lower-level interpolations.

Morrison's central argument. The score of *Vers le blanc* shows only the pitch interpolation. The archival code shows that Saariaho was simultaneously interpolating timbre, rhythm, vibrato, tremolo, and spatial phoneme distributions. Callender's analysis, though mathematically profound, is therefore an analysis of the *work-as-seen* rather than the *work-as-heard*. The true continuous transformation is multidimensional and cannot be reduced to voice-leading geometry alone.

8.5 How the Code Implements This Vision

Veir le Blaunk is a direct computational synthesis of both perspectives:

Layer	Callender/Saariaho original	This code
Pitch	Linear 3-voice glide	4-voice mystic tetrad chain
Timbre	CHANT phoneme interpolation	4-formant vowel-morph
Rhythm	Amplitude modulation patterns	4-channel mute-gate generator
Space	4-channel diffusion	19 stereo-pair output
Structure	Palindromic macro-form	Spiral IOI evolution

§9 Scriabin's Mystic Tetrachords in Context

9.1 From Hexachord to Tetrachord

Scriabin's famous "mystic chord" is conventionally rendered as a hexachord:

$$\{0, 6, 10, 16, 21, 26\} \equiv \{0, 6, 10, 4, 9, 2\} \pmod{12} \tag{16}$$

which in C major spelling is $C - F\sharp - B\flat - E - A - D$.

The code *deliberately truncates* this to a **tetrachord**:

$$\boxed{[0, 6, 10, 16]} \tag{17}$$

In C: $\{C, F\sharp, B\flat, E\}$. This retains the most harmonically characteristic intervals of the mystic sound—the tritone (0,6) and the minor seventh (0,10)—while discarding the outer voices that would anchor it too clearly to a single tonal orientation.

9.2 Interval Structure and Set-Class Identity

The interval vector of the tetrachord $[0, 6, 10, 16] \pmod{12}$:

Interval class	1	2	3	4	5	6
Count	0	1	1	2	1	1

Mod-12: $\{0, 6, 10, 4\}$, sorted: $\{0, 4, 6, 10\}$ = set class $[\mathbf{0258}]$. This is the *French augmented-sixth tetrachord*, closely related to the all-combinatorial hexachord territory of the mystic chord, that can also be heard in Scriabin's 'Vers La Flamme'.

9.3 The Cluster Sequence == Voice-Leading Path

The eleven clusters in the code span the following root pitches (in MIDI):

$$r = \{48, 42, 50, 45, 53, 47, 55, 40, 52, 44, 48\} \tag{18}$$

The motion between successive roots follows an irregular intervallic profile:$\{-6, +8, -5, +8, -6, +8, -15, +12, -8, +4\}$ (in semitones). These are not random: the alternating descents and ascents of around 5–8 semitones trace a *zigzag path through the circle of fifths neighbourhood*, ensuring that successive chords share no common tones (maximising harmonic novelty) while remaining within a register that keeps all four voices in a comfortable mid-range.

The opening and closing chords use the `/4` and `/10` scalings:

$$\text{Opening}: \quad \{48, 42, 50, 45, \ldots\}/4 \;\Rightarrow\; \text{very low frequencie} \tag{19}$$

$$\text{Closing}: \quad \{48, \ldots\}/10 \;\Rightarrow\; \text{extremely low, near-inaudible sub-sonic register} \tag{20}$$

9.4 Placement in Callender's T-class Space

Each mystic tetrachord $[0, 6, 10, 16]$ built on root r maps to a point in the 3-dimensional T-class space for four-voice sets. The trajectory through 11 such points over 600 seconds is a *piecewise linear* path in this space, with a curve parameter of -4 causing it to decelerate into each target chord before moving on. The path passes through multiple inversional and permutational boundaries, generating the complex reflective structure that Callender describes for *Vers le blanc* — but now at a larger scale, traversing a full ten-segment trajectory.

§10 Spiral Time Structuring: Theoretical Basis

10.1 The Spiral as a Compositional Metaphor

Morrison's account of the Saariaho archives notes that for *Jardin secret II* (1985–86), Saariaho used a *spiral diagram* to map the work's large-scale structure: a series of cyclic passes through five types of material, each revolution taking longer than the last, enacting a proportional augmentation of sectional durations. A similar template appears in *Nymphéa* (1987). The spiral metaphor captures a fundamental compositional desire: to create form that is simultaneously *repetitive* (same material cycles back) and *progressive* (each return is expanded in time, and thus transformed in character).

10.2 Formal Definition of Spiral IOI Evolution

The `\spiral` mode in `~generateMutePatterns` implements this idea within the domain of rhythmic gate patterns. Given:

- a base pattern of $N_h \in [6, 14]$ inter-onset intervals $\{d_j\}_{j=1}^{N_h}$ drawn from $\mathrm{Exp}(0.3, 2.0)$ and normalised to segment length L;
- $R \in [3, 8]$ repetitions of this pattern;
- a target compression/expansion factor $\delta_{\text{target}} \in [0.5, 0.8]$ (shorten) or $[1.2, 1.6]$ (lengthen);

the evolved IOI for repetition r and beat j is:

$$d_j^{(r)} = \frac{\delta(r) \cdot d_j^{(0)}}{\sum_{j'} \delta(r) \cdot d_{j'}^{(0)}} \cdot L, \qquad \delta(r) = 1 + \frac{r}{R-1}\big(\delta_{\text{target}} - 1\big) \tag{21}$$

Because the denominator renormalises to L, *the total pattern duration never changes* — only the internal distribution of time shifts. In a shortening spiral ($\delta_{\text{target}} < 1$), longer gaps compress faster than shorter ones, causing the rhythm to concentrate into brief bursts separated by increasing silences. In a lengthening spiral, the converse occurs: brief gestures expand into longer events.

10.3 Connection to Nancarrow's Acceleration Canons

Callender's paper extends the continuous-transformation framework to *tempo interpolation*, analysing Nancarrow's Study No. 22 as a trajectory in tempo T-class space. The three voices accelerate at rates 1%, 1.5%, and 2.25% respectively. The temporal convergences (where two tempos momentarily coincide) appear as reflections off the boundaries of the normal half-region, structurally analogous to the voice crossings in *Vers le blanc*.

The spiral IOI evolution in this code is a discrete-domain approximation of Nancarrow-style acceleration applied to the *gating structure* of formant channels. Rather than continuous tempo change, it produces a *proportional reweighting* of rhythmic events — the same sense of time stretching or compressing within a fixed duration. Crucially, this operates independently for each of the four formant channels, so that at any given moment, different spectral regions of the sound may be pulsing at different evolved rates (a four-voice tempo canon enacted within the timbral domain).

10.4 Interaction with Four Formant Oscillators

The combination of spiral time structuring with four *independently gated* formant channels produces what might be termed a *spectro-rhythmic counterpoint*:

1. At any moment, $0, 1, 2, 3$, or 4 formant bands are audible.
2. When fewer formants are active, the timbre shifts — the vowel character of the sound changes, because the perceived vowel identity depends on the *relative presence* of the four formant regions. With only F1 active, the sound resembles an open vowel (a, o). With only F3 active, it resembles a high front vowel (i). The mute pattern thus controls timbral identity through selective spectral suppression.

3. The spiral evolution means that *over time*, the rhythmic character of each formant's envelope gradually accelerates or decelerates, causing macro-scale timbral rhythms to emerge and fade.

This is the computational realisation of what Saariaho describes as a "multi-dimensional network in which detail is strictly controlled on several levels."

§11 Performance on the GRM Acousmonium

11.1 What Is the Acousmonium?

The *acousmonium* (coined by François Bayle at Groupe de Recherches Musicales, Paris) is a large-format loudspeaker orchestra designed for the projection and spatial diffusion of electroacoustic music in concert. The GRM's own acousmonium comprises between 50 and 80 loudspeakers of diverse sizes, directional characteristics, and frequency responses, arranged in a semicircular or full-surround configuration around the audience. A performer (the *speaker pilot*) controls a mixing console to route and spatialise the sound in real time, treating the loudspeaker array as an instrument.

The aesthetic of acousmonium performance is inseparable from the *acousmatic* tradition initiated by Pierre Schaeffer: sound is emancipated from its source, experienced as pure sonic phenomenon in space. As Pierre Henry put it, the loudspeaker is "the instrument, and space is the score."

The incommensurate LFPar modulation rates (125 s and 150 s) produce a spatial trajectory with a period of $\mathrm{lcm}(125, 150) = 750\,\mathrm{s}$ — longer than the entire piece, so the spatial movement never literally repeats.

11.2 Relationship to Bayle's Acousmatic Aesthetic

Veir le Blaunk instantiates not one but several overlapping i-sons (*image sonores*):

- The *human voice* — suggested by the formant structure but never cleanly realised
- The *chord* — the mystic tetrachords hover at the threshold of harmonic recognition, too slowly evolving to be "heard as a series of changing chords" (Saariaho's own phrase).
- The *breath* — the spiral mute patterns

§12 Synthèse

Bringing together the preceding analyses, the following table summarises the multi-dimensional network of continuously interpolating parameters in the piece:

Parameter	Start state	End state	Duration
Pitch register	Sub-bass (MIDI/4)	Sub-sonic (MIDI/10)	600 s
Harmonic content	Mystic tetrachord on C4	Mystic tetrachord on C4	600 s
Vowel sequence	`\rando` (noise)	`\aw` (open)	600 s
Formant F1 gate	Depends on section mode	Depends on section mode	10×60 s
Tremolo rate	∼6.25 Hz	Slowly drifting	600 s
Spatial width	0.5–1.0 (LFPar)	Cycling (750 s period)	600 s
Spiral IOI	Base pattern	Compressed/expanded	Per section

Each parameter evolves on its own timescale, in the manner of Saariaho's "multiparametric network." No two parameters share the same period or phase, ensuring that the composite texture never exactly repeats.

§13 Appendix: Key Mathematical Summary

13.1 Callender's T-class Distance (General Form)

$$\rho_n(A,B) = \sqrt{\frac{n}{n-1}\left[\sum_{i=1}^{n}(b_i-a_i)^2 - \frac{1}{n}\left(\sum_{i=1}^{n}(b_i-a_i)\right)^2\right]} \tag{22}$$

13.2 Spiral IOI Evolution

$$d_j^{(r)} = \frac{d_j^{(0)}\cdot\delta(r)}{\sum_{j'} d_{j'}^{(0)}\cdot\delta(r)}\cdot L, \qquad \delta(r) = 1 + \frac{r}{R-1}(\delta_{\mathrm{T}}-1), \quad r = 0,\dots,R-1 \tag{23}$$

13.3 Formant Interpolation

$$F_k(t) = \exp\Big[(1-\alpha(t))\ln F_k^{(n)} + \alpha(t)\ln F_k^{(n+1)}\Big] \tag{24}$$

where $\alpha(t) = \big\{(t/T)(N-1)\big\}$ is the fractional position within the current vowel segment.

13.4 Mystic Tetrachord T-class Position

The interval vector of $[0,6,10,16]$ (mod 12) places the set at coordinates $\langle 6,10\rangle_T$ in Callender's oblique T-class plane. Its inversion $[0,2,6,12]$ (mod 12) is reflected about the line $v_3 - v_2 = v_2 - v_1$ and belongs to the same set class [0258].

Analysis and documentation for ***Veir le Blaunk****, a work for acousmonium and computer-generated sound.*

Theoretical frameworks: Callender (2004), Morrison (2021), Saariaho (1987).