Title
Scenes, For Two-Channel Fixed Media

Author
Rostovtsev, Ilya Yurievich

Publication Date
2017

Scenes, For Two-Channel Fixed Media

By

Ilya Yurievich Rostovtsev

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Music

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Myra Melford, Chair

Professor Edmund Campion

Professor Björn Hartmann

Summer 2017

Abstract

Scenes, For Two-Channel Fixed Media by

Ilya Yurievich Rostovtsev

Doctor of Philosophy in Music

University of California, Berkeley

Professor Myra Melford, Chair

This paper is an account of the creative process that resulted in the dissertation for a PhD in music composition, *Scenes*, a two-channel fixed-media composition of electroacoustic music. The commentary is provided as a supplement to the music composition that forms the primary work carried out for the dissertation.

The practice of studio composition demands from the composer both the conception and realization of a musical work. The effect of interdisciplinary practice necessitated by these demands informs both technical and aesthetic approaches. Software tools written by the composer with a specific goal of realizing an aesthetic intent provide an insight into the discovery of musical language. Conceived to serve the music, the defining features of the toolset outline the nature of the musical problem that the technology was designed to solve.

The composition of *Scenes* is a combination of imagined behaviors, embodied improvisation, and experimentation with tools. Discussion of the novel tools generated for the piece illustrate the aesthetics of *Scenes*. The first, *trajectories*, provides kinetic spatialization effects to static sounds. The second, *utterance*, transforms composer's vocal improvisations into electroacoustic sound. The tools presented here are not mere sound processing devices, readily available for a particular effect in the music of others. Rather, they are a manifestation of the act of music making, of composition, capable of transforming their author's ideas based on the lessons learned from their conception and implementation.

Table of Contents

Elements of the Composition

The practice of studio composition demands from the composer both the conception and realization of a musical work. In addition to my role as a composer, I was also the performer, improviser, engineer, and tools builder—roles necessitated by the act of crafting a musical experience. In contrast to the social, collaborative approach of music making that separates the roles of the work's conceiver from its interpreters and, in case of studio work, audio engineers, electroacoustic composer is an experienced practitioner of all of the above. Experiences obtained from the practice of such an interdisciplinary approach informs all aspects of the resulting music. Tracing a path to the development of ideas heard in the final work is therefore difficult: tools were inspired by musical concepts, experimentation with tools led to new musical ideas, which led to problems of sound design, solutions to which inspired new compositional strategies for structuring musical behaviors. All tasks lead back to the music.

Electroacoustic composition does not fit well within the notion of expressive agency by the composer in realization of her work—a practical realization of an idea rarely provides satisfactory results, requiring edits and reconstruction that often undermines or negates the original intent. Nor does it fit within the notions of improvisatory or purely experimental pieces, realized without complete preconception of the final result—the music is evaluated at every step of the production process due to the recorded medium. The agency of electroacoustic composers often lies in *curation* of experiments carried out in the studio. Conception of new musical behaviors resulting from the experience of sound production relates experimental music-making with compositional intent. My personal relationship with studio work is best defined as an obsessive exploration of electroacoustic sound, acts of experimentation en route to discovery and learning. The piece is a discovery of a musical language with which the compositional structures can manifest.

Regretfully, valuable mistakes that served as stepping stones in the creative process are rarely available to present as evidence. The evolution of a musical concept from early ideas to realized studies to final work is rarely organized as such in the creator's mind. An honest reflection on my own creative practice is likely to conclude that I am unaware of the stage of the composition process until after the musical fragment being worked on is considered finished. But the process is far from subconscious: audition of studio-produced content seems to consistently result in plans, to-do lists, and additional steps in shaping the content into a convincing musical work.

Tools written by the composer with a specific goal of realizing a particular musical project provide an insight into both the aforementioned mistakes and the discovery of the musical language. All of the tools created for *Scenes* are specifically crafted to address a particular set of compositional problems. Since the primary focus was on composition rather than software, the tools are "incomplete" for public consumption, left in the state most beneficial to the composer. The features of software tools are integral to the compositional process and indicative of the musical values for the particular composition. Similarly, the choice to leave features unimplemented is also interesting. In some cases, the work could be done via manual construction—features are omitted for practical reasons. In some cases, the lack of a feature may signal its failure to contribute to the act of composition.

This paper is an account of the creative process that resulted in the dissertation for a PhD in music composition, *Scenes*, a two-channel fixed-media composition of electroacoustic music. To tell the tale of composing the piece, aesthetics of the finished work are summarized first. The discussion of custom tools and their role in the production of the piece is used to reflect on the circuitous paths and discoveries that led to realization of the musical language. *Scenes* was not composed as much as it was discovered from experimentation, improvisation, production, and play.

The Piece

Aesthetics of *Scenes*

Inspired by a programmatic intent of a *ballet of glass*—a dynamic, intricate space populated by fleeting, colliding glass shards, assembling into an unseen structure experienced aurally, *Scenes* is an animated study of glass resonance. The program was the motivation for early assemblies of sound, tools created during the production, and formal considerations of the piece. The practice of manufacturing sounds and their combinations, however, continued to inform the work being constructed. With the passage of time, and changes in focus, the *ballet of glass* became a metaphor for working with the traditional elements of music: rhythmic structures, pitch collections, and contrasts of timbres and textures. Yet, while the compositional focus shifted to manipulating musical structures, both tools and sounds conceived in early stages of the piece carried with them the focus on spatialized gestures, animated behaviors, and timbres of glass resonance as the primary sound material used in the work.

My previous acousmatic work, *Understatements*, was inspired by writings of Russian Constructivists and the notion of *faktura*—the material essence of media used in production.[1] Alexey Gan defined it as "a material knowingly chosen and rationally deployed," a definition best suited for musical application.[2] Marc Battier's proposal of using the term in analysis of electroacoustic music,[3] combined with Dennis Smalley's notion of *spectromorphology*[4] and Natasha Barrett's strategies for spatio-musical composition[5] contributed to the musical thinking informing past work.

Understatements was realized in collaboration with acoustic performers: instrumentalists were invited to the studio, providing timbres, articulations, and behaviors for use in studio production. Digital manipulation of sounds was focused on creating moments of sonic unfamiliarity in music obsessed with material properties of its sources (acoustic instruments). Source bonding, ability of a listener to identify or associate sound with its source, was considered the focus of compositional structure in *Understatements*.

In contrast, *Scenes* utilizes synthesized and generated sounds. In spite of its focus on glass as the source of textural properties and source-bonding associations, the music prioritizes behaviors of sounds, their articulation, rhythmic combinations, and textural evolution. *Scenes* is not obsessed with the acoustic totality of glass, but with the emergent structures of sounds and their articula-

---

1 Ilya Y. Rostovtsev, Source-bonding as a variable in electroacoustic composition: Faktura and acoustics in “Understatements,” University of North Texas (2010).

2 Marc Battier, “What the GRM brought to music,” *Organised Sound*, 12(3) (2007): 193.

3 Marc Battier, “A Constructivist Approach to the Analysis of Electronic Music and Audio Art—between instruments and *faktura*,” *Organized Sound,* 8(3) (2003).

4 Denis Smalley, “Spectromorphology: explaining sound-shapes,” *Organised Sound* 2(2) (1997): 110-111.

5 Natasha Barrett, “Spatio-musical composition strategies,” *Organised Sound,* 7(3) (2002): 315.

tion. Glass objects may be struck, bowed, rubbed, and otherwise made resonant. Yet glass can be thrown, tossed, spun, rolled, crushed, and cracked. The latter, impractical in live acoustic settings, were chosen as the main focus of the electroacoustic production.

Scenes is a work of acousmatic fiction, in which glass resonates by itself, emerges from silence and resonates without being activated by a human agent. The piece is inspired by kinetic explorations of electroacoustic works, Palle Dahlstedt's *Gummi*, Luigi Ceccarelli's *Cadenza Esplosa*, timbral designs of Robert Normandeau's *Murmures* and Erik Michael Karlsson's *Un et Deux*, and noise compositions of Emelie Payeur. Yet *Scenes* was also produced with an explicit desire of eliciting the intimate settings of small chamber ensembles. Works of Salvatore Sciarrino, Giacinto Scelsi, Luciano Berio and solo works of Helmut Lachenmann inspired musical constructions in *Scenes*, its scarce materials, and developmental focus.

Materials

The primary material in *Scenes* comes from twenty two attacks of glass sounds, obtained from an Adobe sample library licensed to UC Berkeley students. Sources of broken or excited glass were cut: short attacks, each under a second in length, were kept. The resulting set of twenty two glass impulses amounts to five seconds of source sound. Resonances were synthesized by prolonging impulses with reverb—an ancient technique of electroacoustic production dating back to early twentieth century works by Pierre Schafer and Pierre Henry. Multiple reverb tails excited by the same impulse were filtered separately and combined together to create a set of seven resonances for each of the twenty two sources. Each of the resulting resonances preserves the original spectra of the impulse but varies the decay and presence of partials, providing timbral variations of impulse source spectra. The synthesis of resonances was not aimed as eliciting a spatial presence of room effects, but rather of a continuing excitement of glass with a sustained *legato* articulation.

Custom softwares used for processing these sounds were *trajectories*, a binaural spatialization tool with doppler shift and stereo imaging, and *utterance*, a set of scripts for applying amplitude envelopes of existing concrete sounds to the resonances outlined above. Concrete material for *utterance* processing came in the form of vocal percussive sounds improvised by the composer. *Scenes* utilizes a number of music production softwares: the piece is arranged in Ableton Live,[6] uses Max/MSP[7] for custom processing effects, synthesizer plugins for generating low-frequency bass sounds, and digital audio effects for typical electronic sound production: compression, EQ, and distortion.[8]

---

[6] "Learn More About Our Music Our Music Making Software Live | Ableton," Ableton, July 20, 2017, https://www.ableton.com/en/live/

[7] "Max Software Tools for Media," Cycling 74, July 20, 2017, https://cycling74.com/products/max

[8] Plugin instruments and effects used in the piece were built-in instruments available in Ableton Live, and plugins from Waves (http://waves.com/) and Native Instruments (https://www.native-instruments.com/en/).

Synthesized resonances are the primary source of sounds for the duration of the piece. They were transformed into spatial gestures, rhythmic behaviors, and vertical aggregates (chord-like constructions that utilize combinations of noise and pitch). New materials were processed anew for additional sounds: reflections upon reflections of the source resonance and its transformations. Evolutions of sound through electroacoustic manipulation led to an exponential multiplicity of sources used in the work.

Scenes is a single-movement work—in spite of presence of sections within the piece, effort was made to make transitions between different sections as seamless as possible. The following time points delineate the section organization of the piece during the composition process, informal names of the sections in parentheses:

1. Introduction (flight): 0:00—3:00
2. Rhythmic development (sparks): 3:00—6:43
3. Resonance (chime): 6:43—8:49
4. Finale (breakage): 8:49—12:24

The tools created for the piece are responsible for the salient voices in all but the third section. The introduction explores articulation of glass resonances through spatial manipulation, using the *trajectories* software to simulate movement of sounds. The second and fourth sections of the piece are constructed using transformed vocal improvisations of rhythmic behaviors. The third section is composed entirely out of the original set of resonant materials, arranged into spectral aggregates spanning the entirety of human hearing range. The third section manipulates source resonances with traditional effects (distortion, EQ, tremolo, and reverb).

The tools created for realizing the piece are a part of its composition. Their primary evaluation is in enabling sound design practices to carry out the creative work. Yet, more importantly, the act of implementing the tools informed the musical language of the piece. For example, sections that do not utilize *trajectories* in their sound design nonetheless incorporate manual manipulation of pitch and amplitude, inspired by the shapes observed during authoring, debugging, and testing the tool. Such "mimicry" of algorithmic sound shaping, in turn, inspired new ideas for working with the tools.

The discussion of tools utilized in the piece is not limited to mere technical details of the softwares created for *Scenes*. Rather, the account is focused on the aesthetics of music making, on musical qualities of the softwares, and their relationship to the compositional ideas.

The Tools

Trajectories

A Moving Sound

The opening section of *Scenes* explores the sonic illusion of motion. Observing a sound source moving through space alters the perception of its loudness (depending on proximity of sound source to the observer), pitch (doppler effect based on the relative velocity toward or away from the observer), and stereoscopic perception of sound. Composing changes of these parameters enables electroacoustic composers to create compelling illusions of moving, flying, fleeting, and spinning sounds. This effect is difficult to author using manual manipulations of pitch, amplitude, and stereo positions. *Trajectories* is a collection of softwares providing (1) a graphical user interface for authoring motion paths, and (2) a set of DSP routines for applying simulation of motion to sounds.

To illustrate the musical meaning of spatial movement, consider a simple trajectory of sound motion in Figure 1.[9] The listener is denoted by a head outline in the dashed circle. The movement of a sound source from left to right along the path at some uniform velocity results in:

- a change in dynamics: *crescendo* (as the sound source gets closer to the listener) followed by *diminuendo* (as the sound source passes the listener and moves away);

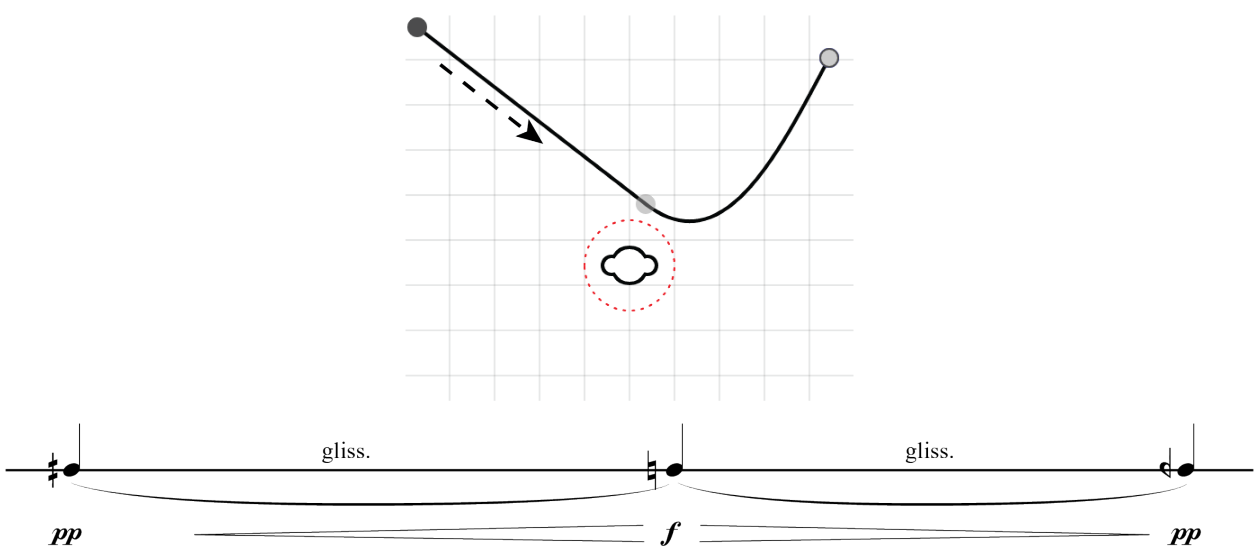

Figure 1: sound source moving along the path (top) and its musical meaning (bottom).

[9] Music notation uses unmetered, proportional notation to estimate the temporal behavior of the resulting phrase.

- a change in pitch: the initial movement of sound towards the listener results in a sharper sound, motion away results in a lowered perceived pitch (actual tuning depends on the velocity of the moving sound source);

- a change in stereoscopic perception of sound origin from left to right.

The musical score with microtonal pitch notation used to illustrate the behavior is an imprecise sketch of the expected behavior. To compute the values of dynamics and pitch, each square of the grid in the path diagram was chosen to represent a distance of 2 meters, making the start point of the sound (black dot) to be, roughly, 13.5 meters away from the listener. The time for traversing the entire path was chosen to be two seconds. In such a case, rough estimate of the velocity is 12 meters per second, based on an estimated length of the path. Results of the computation traversing the path curve of Figure 1 given the above assumptions are shown by the graphs in Figure 2,[10] representing the amplitude changes, doppler shift (as a ratio of the source frequencies), and panning.[11]

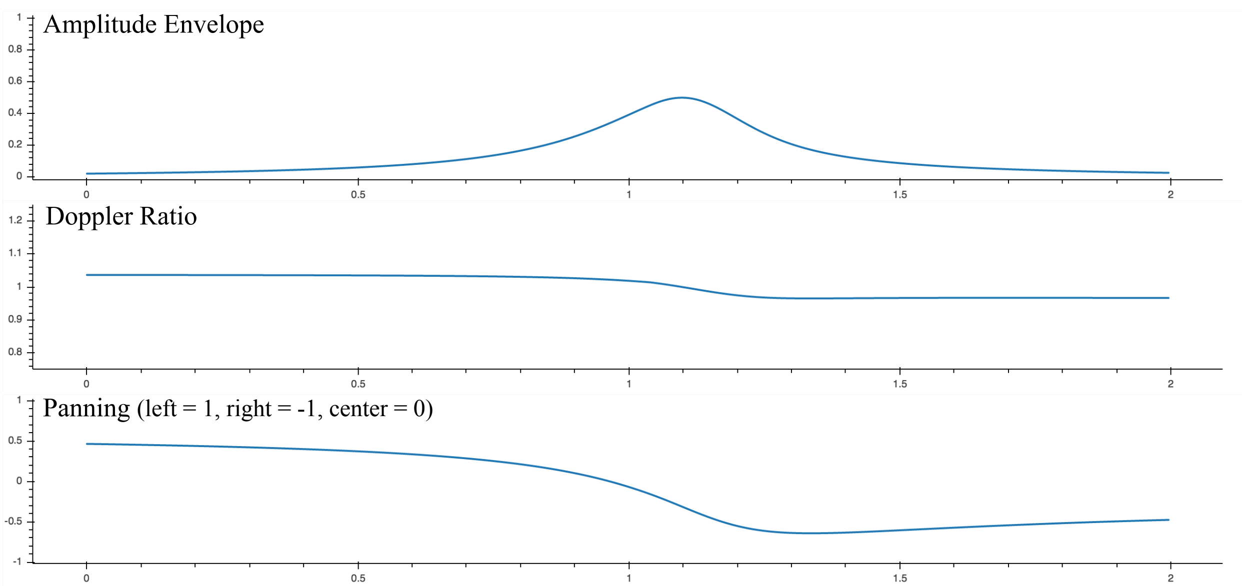

Figure 2: computed results of the path shown in Figure 1.

If the motion is articulated by a spiral, a rhythmic pattern emerges (Figure 3):

- changes in dynamics articulate a set of accents;

---

[10] The graphs of sound data computed by Trajectories are rendered using Bokeh visualization library from Continuum Analytics, http://bokeh.pydata.org/en/latest/

[11] Due to the relative simplicity of the math, and its distraction from the narrative focused on musical behaviors, the formula used to compute the doppler, amplitude, and pan can be found in the appendix.

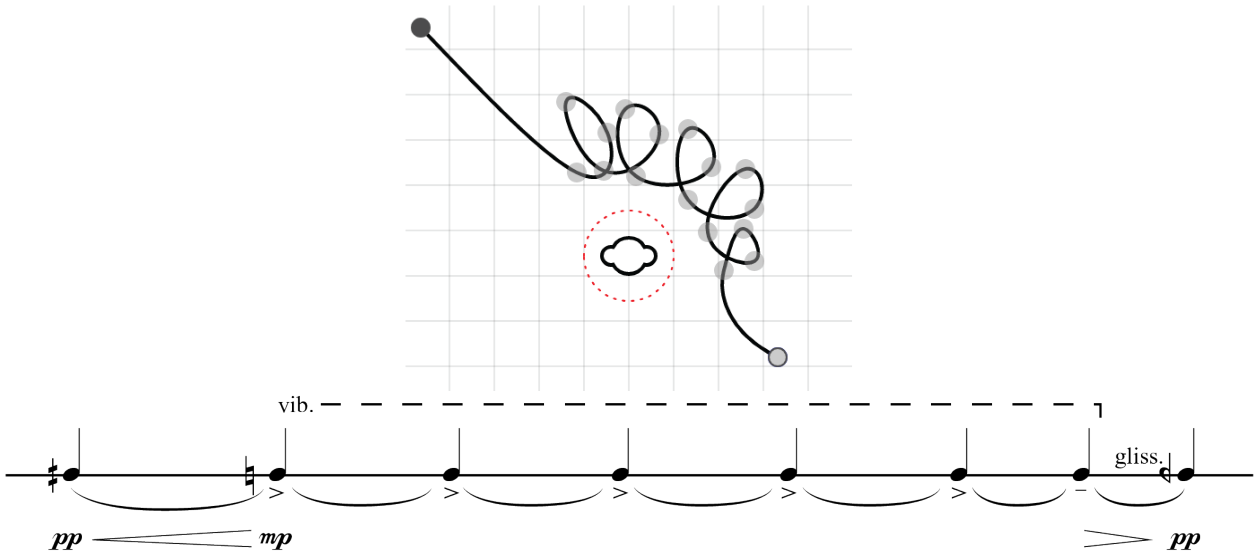

Figure 3: movement articulated by a spiraling motion.

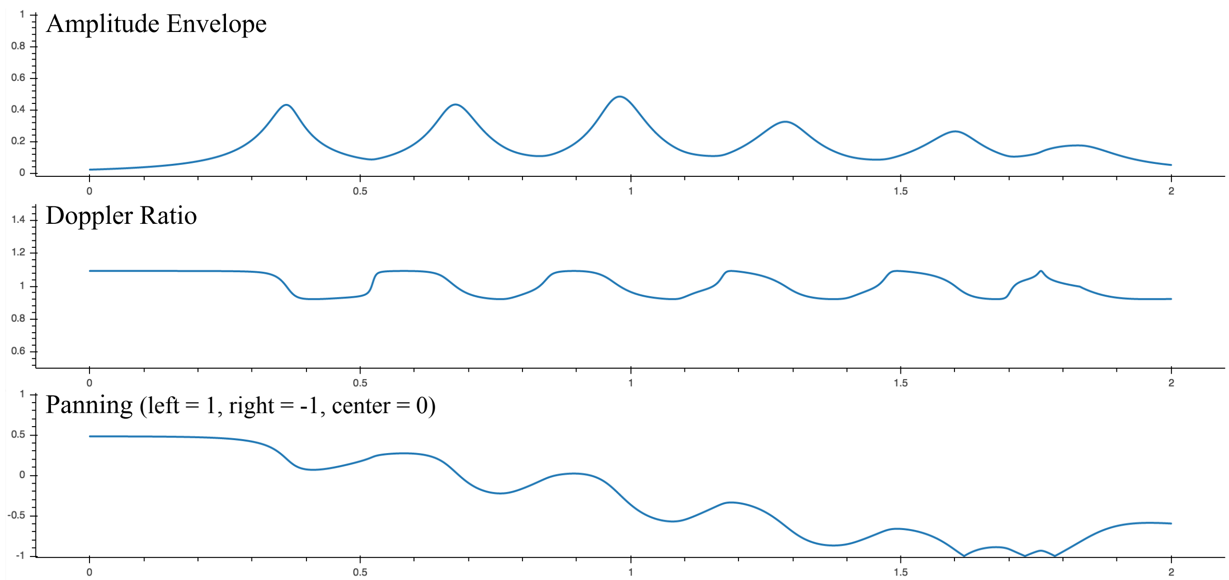

Figure 4a: computed envelopes of the spiraling movement.

- doppler effect provides a *vibrato* effect (pitch moves up and down according to movement of the sound source toward or away from the listener within the spiral);

- the panning movement curve, in spite of its general movement from left to right, becomes more complex—imagining its shape is difficult.

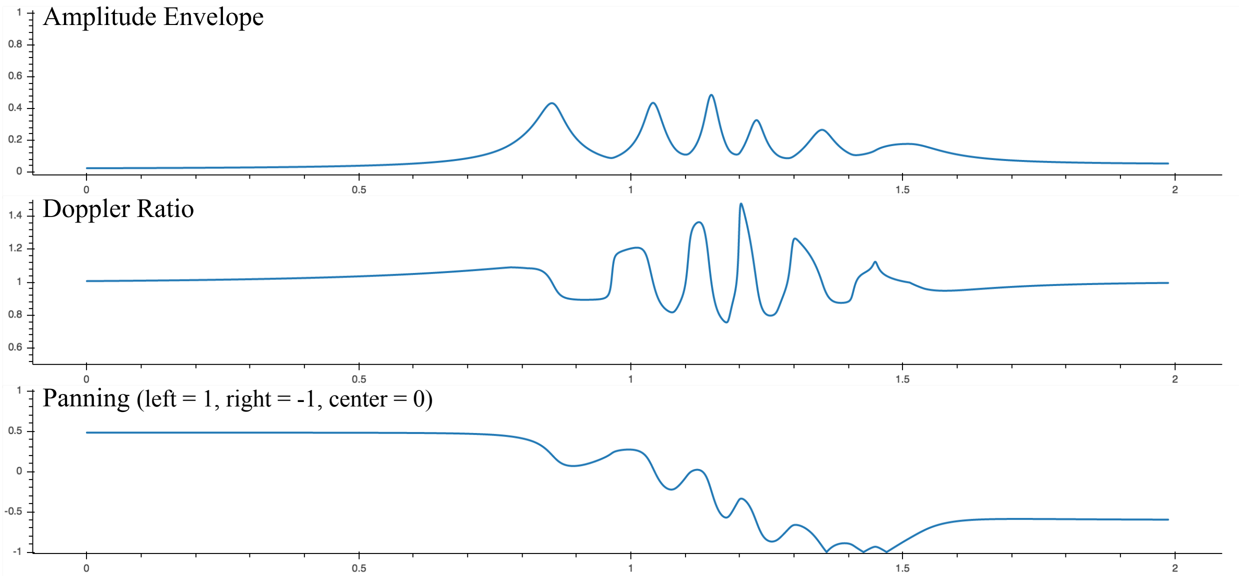

Figure 4b: computed envelopes of spiraling movement with varied motion across the path.

The unfolding of the phrase can be summarized as an introductory *crescendo* (with a sharp pitch) followed by a rhythmic articulation of five accents—one for each of the loops in the spiral—followed by a subtle *tenuto* (as the source enters the final arc of the curve) before *diminuendo*. The timing of these articulations is a function of the total length of the curve seen in Figure 3—velocity of movement is constant. Manipulations of velocity can generate dramatic changes in the perception of the rhythmic articulations: a rapid spin across the spiral yields quick staccato-like accents; a slow traversal of the spiral—waves of crescendi and diminuendi. Figures 4a and 4b contrast the envelopes computed with uniform timing across the path (4a) against the amplitude envelope for the same path with varied velocity: the approach and departure of sound were prolonged, while the spiraling movement rushed (4b).

Figure 4b demonstrably "squeezes" the amplitude fluctuations of figure 4a closer together. This has a noticeable effect not only on the timing of the vibrato articulations in the doppler ratio graph, but also on their magnitude. This is to be expected: faster velocity values alter the perceived pitch of a moving sound more dramatically. The slower approach of the sound, however, results in very little doppler attenuation of the source at the onset—the quarter-sharp used to communicate the initial descent of the pitch seen in Figures 1 and 2 no longer applies. Acceleration toward the spiral now leads to a rising of the pitch from the natural tuning, an opposite effect to that of the unaltered path (see Doppler Ratio in figures 4a and 4b).

These basic examples highlight the behaviors exhibited by moving sounds. The effect of the process is rhythmic articulation of vibrato and tremolo, in conjunction with a directional movement of sound in space. Velocity of movement scales the amount of doppler pitch variation. Distances traversed by moving sources change slopes of amplitude envelopes. Timing controls are crucial for shaping the spatial gestures into compelling rhythmic behaviors, achieved by juxtaposing slower and faster articulations of movement.

Digital Sound Spatialization

Spatialization tools developed for *Scenes* borrow from ambisonic encoding[12] for storage of spatialization data and Vector Based Amplitude Panning (VBAP)[13] for virtual source sound positioning. Spatial perception of sound is implemented using Head Related Transfer Functions (HTRF).[14]

Ambisonics encode the spatial location of the sound using the horizontal angle θ, and elevation angle ϕ. Positioning of a source at the desired location is achieved by distributing the signal over ambisonic components with different gains. Elevation was not considered for computational spatialization of gestures in *Scenes* for two reasons. Firstly, the effect of the spatial movement on the musical function, outlined above, is not affected by elevation. The simpler two-dimensional implementation of sound movement (relying only on the horizontal, *azimuth* angle θ) is easier to author than three-dimensional paths, and more suited for rapid prototyping of musical behaviors from motion trajectories. Secondly, since the function of elevation in music is that of different attenuation of partials, a musically similar behavior could be obtained by manual post-processing of the spatialized gestures. Manual sound design techniques, fictitious from the perspective of simulation, are crucial for employing computed behaviors within a musical context.

Current formats for ambisonic encoding combine the signal source for spatialization with the azimuth and elevation data. *Scenes*' exploration of sound gestures required application of the same sound trajectory to multiple sources. The compositional process necessitated splitting the spatialization data from affected sources. This approach was particularly fruitful in manually mixing various sounds undergoing spatialization for maximal effect (as outlined in the Digital Signal Processing section below).

VBAP computes localization of sounds emerging from defined configurations of virtual speakers. In case of *Scenes*, spatialization is carried out by mapping the direction of a sound source to a ring of 72 equidistant virtual speakers. Due to the uniform distribution of the speakers and two-dimensional approach to spatialization, implementation of VBAP was trivial: a virtual sound source position was computed by panning between two adjacent virtual speakers according to the azimuth angle θ. To generate the stereo image of sound emanating from virtual speakers, HRTF filtering was implemented.

---

[12] Jan C. Schacher, and Philippe Kocher, "Ambisonics spatialization tools for max/msp," *Omni* 500, no. 1 (2006).

[13] Ville Pulkki, "Virtual sound source positioning using vector base amplitude panning," *Journal of the audio engineering society* 45, no. 6 (1997): 456-466.

[14] Tilen Potisk, "Head-Related Transfer Function," In Seminar Ia, Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana, Slovenia (2015).

The tool relies on the MIT collection of Head-Related Transfer Functions.[15] Represented by a collection of 144 impulse responses, obtained by placing microphones into the ears of a Kemar dummy head, the data samples the directions of sound at five degree intervals—two impulse responses representing audition in left and right ears for each of the 72 locations in a 360 degree field. Each impulse response is a short audio file containing silence (to model the time difference between sound arriving to each of the ears), and a finite impulse response representation of a filter (frequency attenuation due to the pinna of the ears). Amplitudes of the impulse responses also provide experimentally obtained levels for directional sound reaching each ear. Real-time convolution, using HISSTools,[16] is used to apply the correct set of impulse responses to the single-channel source sound.

The signal processing chain combines doppler-shift pitch variations, distance-based amplitude variations, and the azimuth angle to compute the final effect of spatialization on sound. In addition to these, stereo pan variations and envelopes of velocity changes are computed. This data provides additional controls for exaggeration of spatial presence used in the sound designs carried out for *Scenes.*

To facilitate creation of spatial gestures, a graphical user interface for path authoring was implemented. The user interface application encodes spatialization gesture data as an audio file interpreted by the digital processing chain. Sound design tasks are carried out by first focusing on the trajectory itself, observing rhythmic and pitch-variant behaviors of spatialization using a reference sound (one of the source resonances described in the Materials section). The final sounds were designed by layering several processed resonant sources together. Each of these sources was modulated by the same motion data to ensure the function of the mixture as a single voice.

The User Interface of *Trajectories*

The graphical user interface of *trajectories* is designed to provide explicit control of both the shape of sound movement as well as the timing of the traversal. The user interface for path authoring and manipulation was written in JavaScript with Electron framework for authoring desktop applications.[17] The interface is shown in Figure 5. The path is drawn on a canvas containing a listener in the center. A virtual sound source moves along the path, from the black dot denoting the start point to the grey dot denoting the end. The dashed circle surrounding the listener indicates the *safe zone*—the amplitude of the sound source is not attenuated inside the circle. Additionally, the safe zone circle is used to control the scale: a draggable number in top left denotes the meter value corresponding to the radius of the circle. Changes of the safe zone radius has no

[15] Bill Gardner, and Keith Martin, "HRTF Measurements of a KEMAR Dummy-Head Microphone", July 20, 2017, http://sound.media.mit.edu/resources/KEMAR.html.

[16] Alexander Harker, and Pierre Alexandre Tremblay, "The HISSTools impulse response toolbox: Convolution for the masses," in *Proceedings of the International Computer Music Conference* (2012): 148-155.

[17] The application uses Paper.js library for vector graphics manipulation, Node-WAV library for sound file generation, and OSC-min for OpenSoundControl implementation. Citations can be found in references.

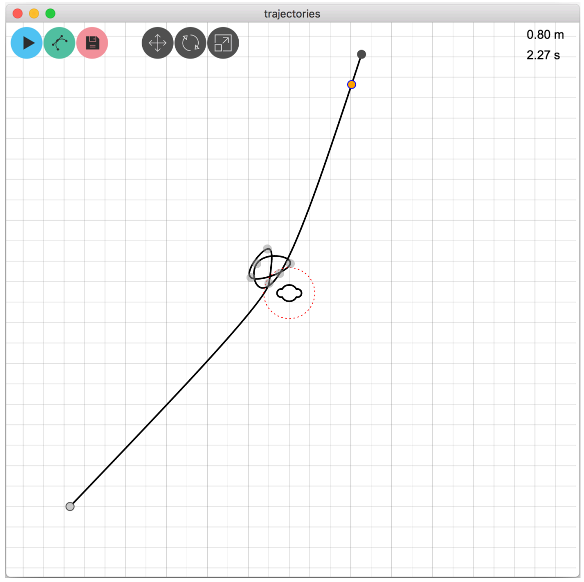

Figure 5: The interface of *trajectories*.

effect on the screen coordinates of the path, but changes the corresponding values of the world coordinates used in simulation of movement, resulting in changes in doppler shift and the perceived volume of the virtual sound source. Below the distance value is a draggable number corresponding to the total time (in seconds) that the virtual sound source would take to traverse the path. The number of seconds is used to compute the default velocity of the virtual sound source along the path.

Icons along the top of the interface allow moving, rotating, and scaling the path with respect to the virtual listener in the center of the canvas. Of these controls, scaling was not used during production, as similar effects were obtained by controlling the safe-zone radius (top-right number) and the timing of the movement. Moving and rotating the path is identical to moving and rotating the listener: the effect is that of observing the motion from a different location in a 2D space. Several of the trajectories used in the piece were recorded with variations in position and rotation of the path to discover new articulation of sounds.

Velocity changes can be made by annotating the path, as illustrated by Figure 6. Dots placed along the path contain a percent value denoting the deviation from the default velocity of the virtual sound source. The grey and white markers placed along the path on either side of the velocity-altering dot are used to denote transitions between the default and augmented velocity. Once the virtual sound source crosses the grey marker, its velocity is linearly interpolated until it reaches the center of the black dot. At the center, the velocity is equal to the average velocity multiplied by the percentage indicated. Similarly, the velocity interpolates back to the default velocity in the time that it takes for the virtual sound source to cross the second (white) marker. Placing the markers closer or further away from the velocity dots manipulates the attack-decay ramps of the velocity change envelope. Multiple velocity annotations are used to create articulated movement along the path. The use of relative values (percentages) in path annotations allows

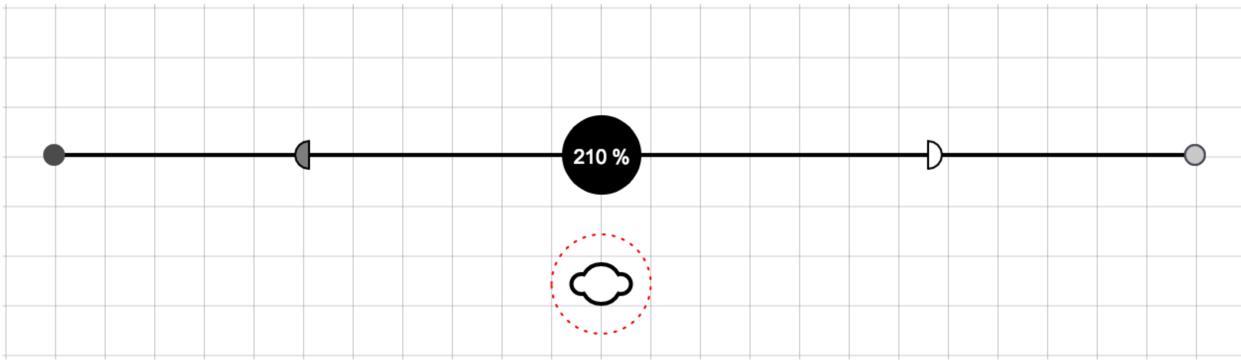

Figure 6: Velocity augmentation indicated by a path annotation.

the time value set in the top-right corner to be used to scale the overall traversal time of the path —percentages apply to the newly computed default velocity.

The JavaScript application has no sound capability. To apply paths to sounds, *trajectories* exports its computation as a 5-channel audio file, containing doppler, distance-based attenuation, azimuth (angle with respect to the sound observer, scaled to the range of floating point values between 0 and 1), panning, and an envelope of velocity changes (Figure 7). The velocity envelope creates an upward ramp during periods of acceleration and a downward ramp for deceleration (zero-values indicate that the sound is traveling at the unaltered default velocity). These envelopes can be used to mix in additional sound sources, e.g. a high-frequency noise band during moments of acceleration to simulate the presence of air. Exported audio file containing motion data is used by the digital signal processing chain, implemented in Max/MSP and detailed in the next subsection. The export procedure computes the results by converting geometry locations (represented by 2D coordinates of the path curve geometry on screen) to world coordinates, a function of the distance value (in meters) indicated by the user. The simulation of motion then runs at a fixed time-step, with the sampling rate 44.1 kHz (CD quality audio)..

To facilitate audition of motion while designing paths, the application also communicates with the Max/MSP sound engine using OpenSoundControl (OSC).[18] The play button seen in the top left corner of the *trajectories* application runs the simulation at an estimated sixty frames per second,[19] sending OSC bundles[20] (key-value stores) containing snapshots of the simulation state to the sound engine. Similar to the audio files exported by the application, the bundles contain doppler effect ratio, amplitude attenuation, azimuth angle, and panning. The velocity curve envelopes are not computed for on-demand path audition and are not supported by the Max patch

[18] Matt Wright, "The Open Sound Control 1.0 Specification," July 20, 2017, http://opensoundcontrol.org/spec-1_0.

[19] The computation in this case does not occur at a fixed time-step assumed when running the simulation for audio file export. The computation during real-time playback simulates the motion using time since the previous frame to achieve correct results.

[20] Adrian Freed, and Andrew Schmeder, "Features and Future of Open Sound Control version 1.1 for NIME," in NIME, vol. 4, no. 06, (2009): 2009.

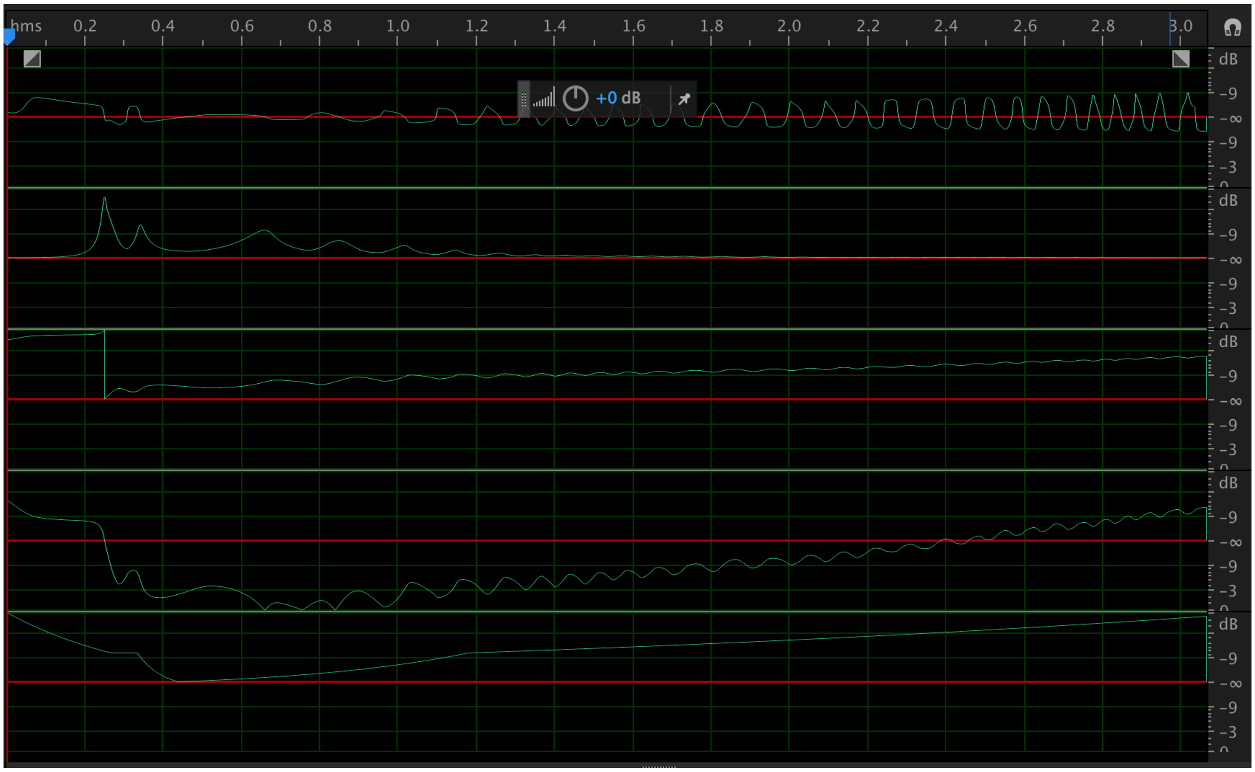

Figure 7: a five channel audio file containing automation data. Top to bottom: doppler shift, amplitude envelope, azimuth angle, panning, and velocity envelope.

used during the process. Layered design using multiple sounds is deferred to working with the renderer using high-resolution audio files containing automation data.

Digital Signal Processing

The five-channel audio file is used by the DSP chain for rendering spatialized sounds. High resolution of control data enables additional fine-tuning of tempo (speeding up or slowing down the playback of automation data) without perceivable side-effects. The signal processing chain is outlined in Figure 8.[21] The sound file produced by trajectories is loaded into *control_data* buffer. The *waveControl* object in the diagram routes the channels of the control audio to the digital signal processing subsystems. From left-to-right, they are:

- doppler, used to control the playback rate of the *groove~*, an object for sound clip playback, containing the source sound being spatialized;
- amplitude, used to multiply the output of *groove~* by the amplitude envelope;

[21] The graphic programming language diagram consists of objects (represented as rectangles) with inputs on top and outputs on the bottom. Dashed yellow cables represent audio data. In this case, control of DSP parameters occurs at audio rate: each audio sample of the *control_data* buffer is used to update the state of DSP parameters.

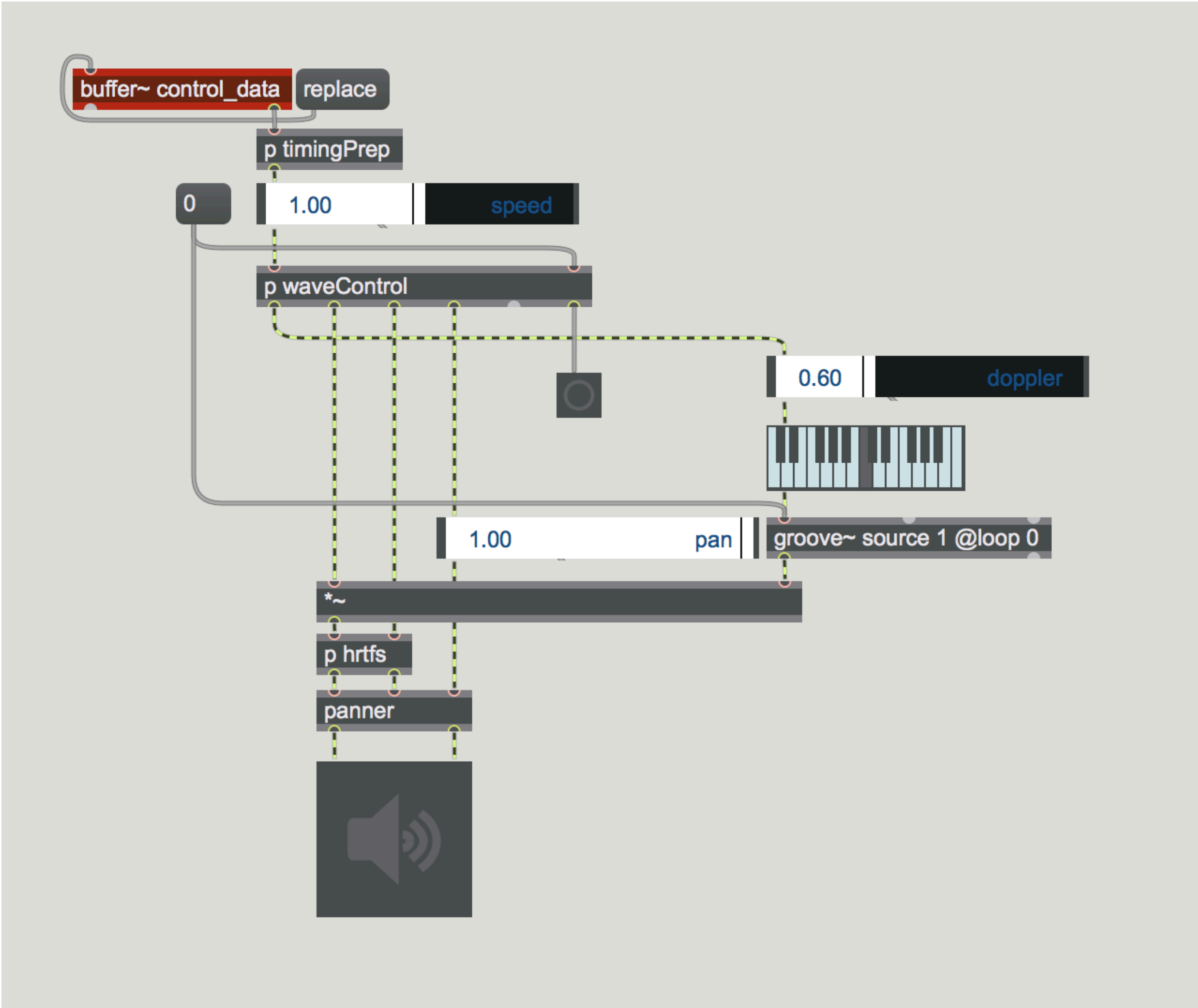

Figure 8: a summary of DSP chain used by the audio renderer as implemented in Max/MSP.

- azimuth, used by the *HTRFs* abstraction, implementing head-related transfer function convolution;
- pan, used for exaggeration of the stereo field;
- velocity envelope, not pictured in this example, is used similarly to the amplitude envelope to control a different layer of sound.

In practice, the above signal chain is duplicated, using the same data to control several layers of sound. The sliders present in the image allow the sound designer to scale the presence of doppler shift and panning. The object resembling a piano keyboard controls the transposition of the source sounds being spatialized. Each sound layer has its own set of scaling sliders to customize the presence of doppler and panning in addition to controlling its pitch.

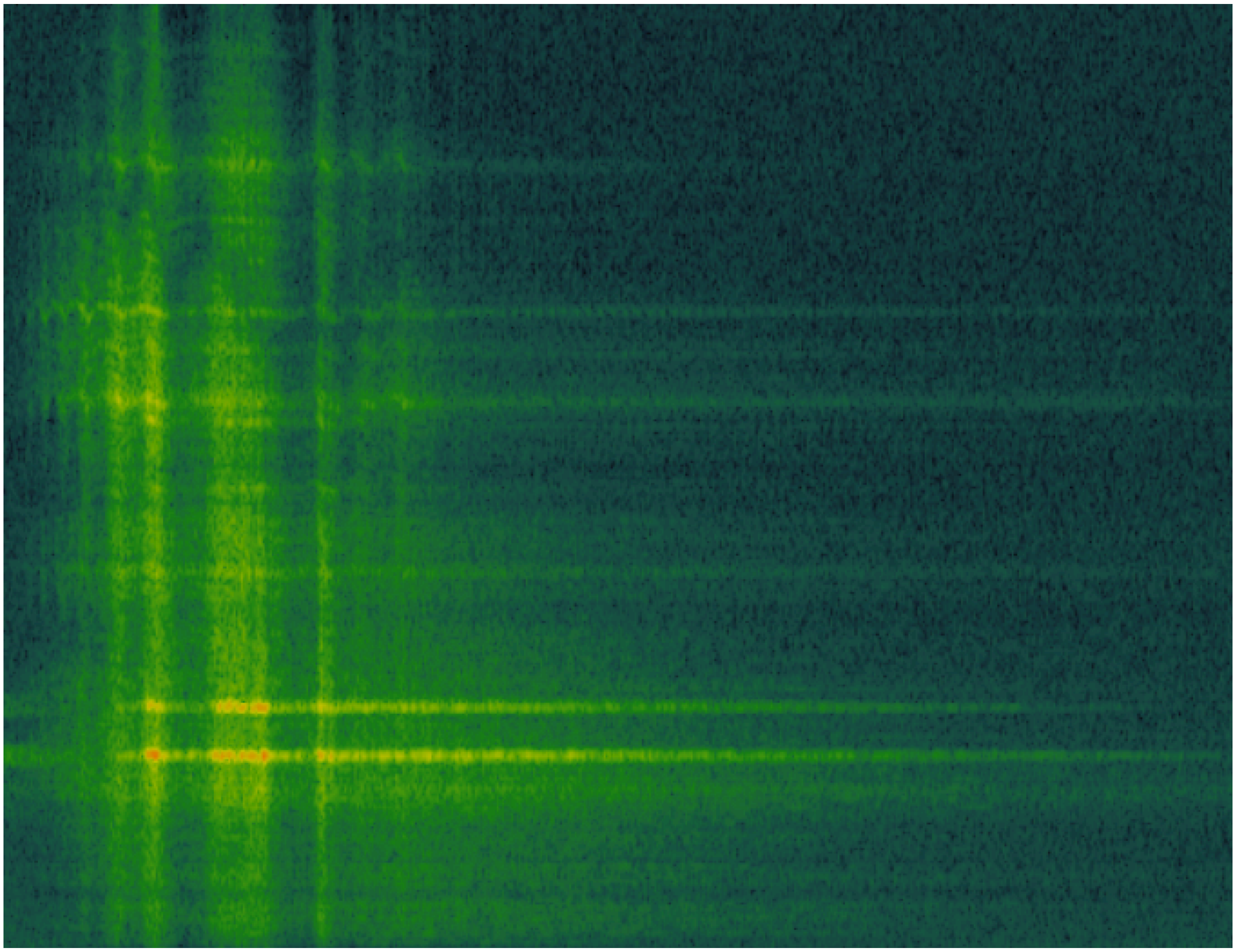

Figure 9: Stable pitched tones (bottom half of the spectrogram) are mixed with doppler-processed noise layers (top part of the spectrogram) illustrating the sound design approach used in the piece. The spectrogram captures frequency spectrum in the 150 Hz (bottom) to 10,000 Hz (top) range. Spectrogram from *Scenes*, at 00:28.

The HTRF abstraction is responsible for sound spatialization approach outlined in the Digital Sound Spatialization section above. Using VBAP, the source is routed to virtual speakers, each corresponding to a different stereo pair of HRTF impulses, and processed accordingly. Amplitude envelope applied to the source sound is precomputed by the *trajectories* application, but is not enough to provide the feeling of distance: post-processing work manually applies reflections and reverberation which provide contextual cues for the proximity of sound in space.

The filtering performed by HRTF transfer functions is most audible with wide-band noise sources. In the case of simple, high-pitch sounds the effect of the filters is a slight alteration of dynamics (sound is slightly quieter when behind the listener). Doppler shift pitch-alterations, computed according to a physical model, provide an additional challenge to the composer: pitched materials constantly glissando, making it difficult to create a sense of a stable harmony or a melodic function. The problems of poor spatialization of high-frequency pitched sounds and unwieldy glissandi were treated as a sound design problem. Pitch sounds that require localization

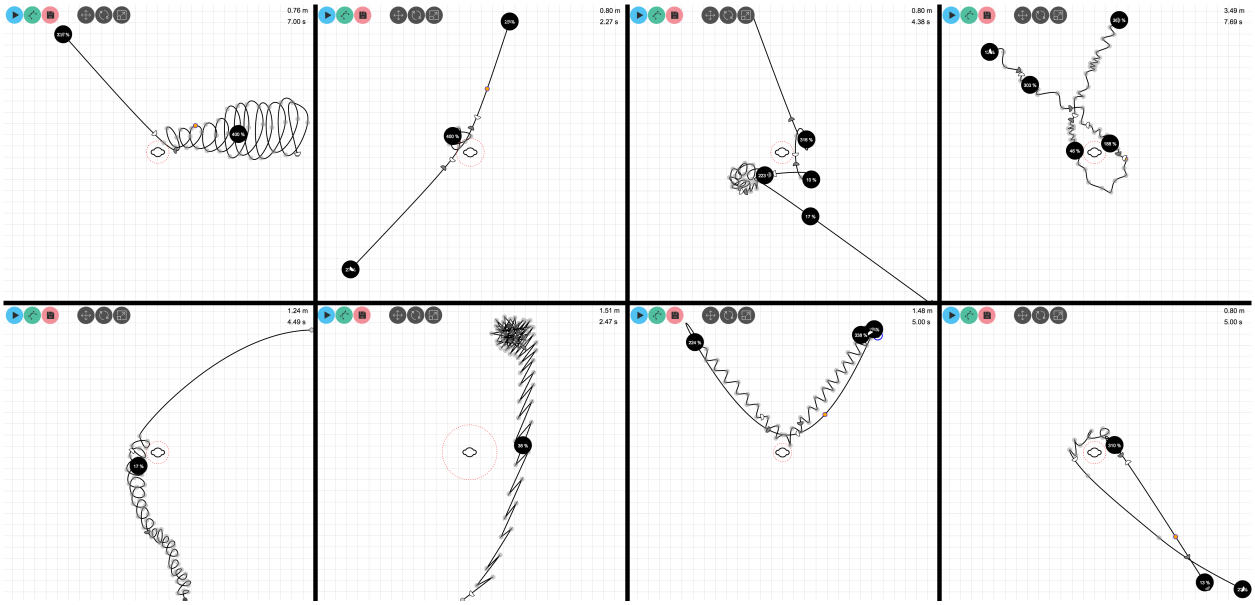

Figure 10: examples of trajectories used in *Scenes.*

(high-frequency inharmonic glass sounds used in the opening three minutes of the composition) were mixed with unpitched noise. The pitched content was left mostly unaltered by doppler shift (enabled by scaling controls added to the Max/MSP rendering engine). The secondary layer of noise (mostly mid-range noise bands between 500 and 8000 Hz), however, is processed with the doppler pitch shift. Due to the lack of a perceivable fundamental frequency of the noise, the effect of the noise glissandi does not interfere with the harmonic implications of pitched sources coinciding with noise sounds. In addition, noise layers provide audible filtering based on the pinna of the ears, improving perceived spatialization. To ensure that noise supported the pitch without being perceived as a separate voice, the layers and synchronized to share common fate—signal-rate implementation of control-rate envelopes ensures deterministic sample-accurate timing across both doppler noise layers and stable pitch sources. Similarly, the panning data was used to amplify the spatialization of sounds least affected by HRTF convolution. The final sounds used in the piece were obtained by manually mixing noise and pitched sounds in a DAW to find the quietest levels of noise layers that nonetheless provided cues of spatialization and movement behavior. The resulting behavior is illustrated by the spectrogram in Figure 9—noise bands modulate in pitch while the pitched layer remains static.

Composing with Trajectories

The sound sources processed by the trajectories path data were rhythmically static—prolonged resonance sounds obtained from samples of shattered glass. The musical goal of sound processing was to create rhythmic articulations with dynamics, vibrato, and tremolo. The opening section utilizing the sounds is a study of rhythmic articulations of high-register pitch content.

The first three minutes of the piece focus on the interplay of a small collection of sounds, each spatialized using the same set of kinetic gestures. Some of the gestures created with *trajectories* are shown in Figure 10. The aural effect of the processing does not result in perception of exact positions or contours seen in the figure: the timings of attacks, the presence of sustain, articulated tremolo and moments of vibrato form the primary musical surface. Additional sounds (sustained glass resonance and low bass sounds) function in counterpoint to the animated activity described above.

Directional awareness provides an additional function in conception of counterpoint, providing relative spatial locations of different voices, even in moments when the same pitch is articulated at once by multiple gestures (e.g. 1:36–1:41). The music is assembled as a set of melodic articulations, rather than a showcase of a physically modeled effect. That is, the aesthetic goal was not that of sequencing flying sounds, but of the music emerging from the flight-like gestures: manual override of physical modeling computation for aesthetic means was necessary to realizing the latter.

Improvised Structures: *Utterance* and Embodied Sound

Improvisation

Composers often summarize complex behavior of sound with concise notation. Electronic music and electroacoustic production requires the composer to realize her rhythmic designs via a manual assembly of sounds files, MIDI events, or parametric automation curves. Both articulation and rhythm require a large set of automated parameters—without them, electronic sounds are lifeless, both in the static behavior of frequency spectra and the machine-driven grid-time of event schedulers.

Electronic music improvisation is an important topic in computer music literature. Research carried out by the New Interfaces for Musical Expression (NIME) community is dedicated to exploration of novel controllers for human-computer interaction.[22] David Wessel's SLABS, a physical instrument designed for improvised performance using electroacoustic sound highlights the importance of real-time control for expressivity. SLABS uses a grid of pressure-sensitive XY-controllers and encodes descriptors of interaction data as audio signals.[23]

Immediacy of sketching sonic figures using one's body is a powerful tool in compositional practice. Contrary to the process of electroacoustic production, vocal and instrumental improvisations are instantly produced. Evaluation may happen instantaneously (via perception of the musical performance), or through recording and analysis. Improvisations are used for schematic reductions of the timings used in the work, and to determine articulations and phrasing of musical passages. For studio work, rather than building a novel controller for improvisation with computer-generated sound, it is feasible to empower composers with methods of translating captured audio improvisations (recordings of vocal or instrumental performance) to electroacoustic textures. Such an approach eliminates the need of editing MIDI or manipulating the timings of sound files: the improvisation is a functional score for generating electroacoustic sound.

The vocal sessions for *Scenes* explored various rhythms, articulations, and vocal mimicry of sonic behaviors already present in the piece—of sounds generated using *trajectories* and of subtle rhythmic features of the source resonances themselves. Recorded passages were originally treated as a source for compositional analysis, extracting rhythmic timings from vocal articulations. During the editing process, many of the sounds perceived as erroneous during recording, and many sounds that I was not aware of producing (odd pauses necessary for breath, recorded exhalations) were found to be of musical value. For example, multiple takes of perceived mistakes during improvisation resulted in unpredictable reiterations of rhythmic fragments (see 3:19—3:28 in the piece). The edits of the vocal sessions resulted in a large collection of rhythmic

---

[22] "NIME | International Conference on New Interfaces for Musical Expression", NIME, July 20, 2017, http://www.nime.org.

[23] Adrian Freed, "Novel and Forgotten Current-steering Techniques for Resistive Multitouch, Duotouch, and Polytouch Position Sensing with Pressure," in *NIME* (2009).

phrases, used to compose a sketch of the second section. The timing of the utterances was not altered, though combinations of concurrent vocal rhythms provided additional rhythmic variety (6:26—6:40).

Digital Mapping

Automating the process of vocal sketching of rhythmic behaviors became desired as large amounts of vocal recordings required substantial manual labor to reproduce electroacoustically. Since the reductions only contained rhythmic material without pitch considerations, experimentation and composition necessary to transform rhythmic reduction into music would require multiple reconstructions of the same rhythmic phrase to achieve the final goal. A set of Python[24] scripts were authored to enable the transformation of vocal utterances into electroacoustic textures. Given a recorded vocal improvisation of rhythm:

1. an amplitude envelope tracing the rhythmic shape of the sound was computed;
2. onset detection delineated attacks in the envelope as separate events;
3. onset data was used to create a textural *collage*, by automatically splicing segments of resonance sources between pairs of detected onsets;
4. collages modulated by the amplitude envelope resulted in rhythmic improvisation gestures rendered with glass textures.

Envelope extraction of the vocal improvisation was achieved with a non-symmetric lowpass filtering of vocal recordings, using the Essentia library.[25] Several approaches to onset detection were considered. Due to the focused, high quality recordings of the source material, detection (novelty) functions used for onset extraction from noisy polyphonic signals were unnecessary.[26] A simple approach to detecting attacks in the signal is to subtract some small value ϵ from the computed amplitude envelope and detect zero-crossings from negative to positive values. The local minimum of the envelope preceding the zero-crossing would effectively locate the actual sample-position of the offset. Although this method locates instances of attack-decay delineations in a computed envelope, it may incorrectly identify waves of amplitude changes in

---

[24] Python Core Team, "Python: A dynamic, open source programming language," Python Software Foundation, July 20, 2017, https://www.python.org/.

[25] Dmitry Bogdanov, Nicolas Wack, Emilia Gómez, Sankalp Gulati, Perfecto Herrera, Oscar Mayor, Gerard Roma, Justin Salamon, José R. Zapata, and Xavier Serra. "Essentia: An Audio Analysis Library for Music Information Retrieval." In *ISMIR*, (2013): 493-498.

[26] Juan Pablo Bello, Laurent Daudet, Samer Abdallah, Chris Duxbury, Mike Davies, and Mark B. Sandler. "A tutorial on onset detection in music signals." *IEEE Transactions on speech and audio processing* 13, no. 5 (2005): 1036.

source sound as new onsets, if minima thereof occurs below the chosen ϵ-value. In this case, the choice of the epsilon value acts as a threshold—signals with amplitude below epsilon would be effectively ignored.

Glass textures are inserted at times corresponding to onset events. The approach outlined above ensures that every attack-decay envelope in an improvised rhythmic source is assigned a new texture. Controlling texture selection is critical to generate musical motion behaving as a single voice, a composition goal for the section. The strategy for ensuring that the process did not result in random, senseless constructions was to utilize textural variations of the same resonance (see Materials) in the collage-generating process. The result is that of random variations of timbre within the expected envelope and timing of the vocal rhythm improvisation.

Vocal improvisations used in the piece frequently reiterated variations on rhythmic articulation of the same sounds. In such cases, timbral effects of the collage distorted the perception of a rhythmic contour—a uniform timbre applied to rhythmic patterns gave far more compelling musical results. Rather than implementing a computation of sound similarity, the concrete behaviors of rhythmic vocal recordings were considered. The presence of strong onsets signified beginnings of musical phrases. Weak onsets and smooth envelopes, categorized by a gradual change in sound level, were considered to be extensions of the phrase rather than delineations requiring new textural material. To implement these considerations, a different approach to onset detection was required. Computed envelope extracted from the improvised rhythm was traversed one sample at a time. If the amplitude change of the envelope from one sample to the next was greater than an experimentally obtained threshold value, an occurrence of an onset was assumed. To avoid subsequent detections in the slope of the same attack, and multiple, rapid onsets resulting in rapid modulation of amplitude in an envelope, every successful detection of an onset “skipped ahead” in its processing of an envelope (by advancing the sample index being used for detection). This approach eliminated instances of rapid onsets.[27]

Onsets are represented by time locations (in samples) in the source audio file. To verify onset-detection and collage-generating procedures, click-track audio files (with clicks corresponding to detected onsets) were generated. To generate sharp transients for attacks, a similar process was utilized, replacing clicks with short attacks[28] of glass sounds, using the original twenty two attacks used to synthesize materials of the piece. These purely percussive sounds became material for the last section of the piece (8:48—12:24).

The scripts were used as an automation of a tedious process, driven by the desire to render rapidly improvised vocal utterances with electroacoustic sounds. The challenge of realizing the second section of the piece (3:00—6:43) was compositional, as the tooling did not account for choices of pitch, timbre, and formal transformations of content.

---

[27] Actual time values for forward skipping of envelope processing were obtained experimentally during processing of sounds, and varied between 100—1000 milliseconds, depending on the content of vocal gestures.

[28] For transient generation, quarter-to-half second edits of breaking glass sounds were used. For sounds of glass cracking used in the fourth section of the piece, onsets were rendered with 50—100 millisecond grains of the same.

Composition

The transformation of material from vocal recordings to glass sounds required choices of pitch and timbre to give the section a musical form. Resonant materials were combined into chords consisting of both noise and pitched sounds transposed and mixed to create new timbres. These chords were either (a) harmonic constructions containing mostly pitched sounds meant to sound together; (b) collections of pitches for melodic construction; or (c) aggregates of noise sounds, which may either sound together or organized in succession. Due to the inharmonic spectra present in source sounds, harmonic constructions were tuned with microtonal adjustments, according to the perceived beating of partials, interactions of noise and pitch sounds, or separation of resonant bands in noise signals. The process is best described as intuitive, embodied evaluations to sound combinations.

The composed timbral progression associates each of the section's rhythmic vocal sources with a collection of transposed resonances. Envelopes of vocal phrases were applied to each component of the corresponding chord. The final sounds were obtained by manual cutting, muting, mixing, and editing the synchronized glass rhythms (corresponding to component sounds of chords) into melodic, harmonic, or noise-articulated behaviors. The musical language of the section was shaped by several processes. The shift from pitch to noise in rhythmic articulations of the piece was a result of the composed timbral progression: chords move from pitch to noise. The mixing and cutting up of the rendered glass rhythms allowed to fine-tune the density of sounds as well as contrast between monophonic (lateral, melodic) and polyphonic (concurrent, harmonic) behaviors in voices. Post processing was used to enhance transients, apply filters, and extend unique events with reverberation.

The rhythmic content of the concluding section of the piece used detected onsets of vocal improvisaions to sequence passages of glass attacks. The musical quality is that of sparse, unpitched rhythmic articulations. Digital distortion of these sounds evoked the sound of cracking glass (8:48—8:51). Due the absence of amplitude envelopes in this case, the original vocal improvisations are difficult to perceive—the sounds function as found rhythms. The concluding section uses the onsets of attacks to delineate changes in animated textures, variations in the mixture of noise and pitch, and the presence of bass sounds.

The compositional strategy to timbral modifications employed in the closing section of the piece is inspired by the collage-generating mechanism implemented for the sounds used in section two. The delineation of timbres using rhythmic events is similar to the computer-controlled collage generation. In the case of the closing section, the process occurs slowly. Rather than providing textural shifts for each attack-decay articulation in an envelope, the last section's timbral manipulations are delineated by longer rhythmic phrases. The source of manipulations are harmonic materials introduced in the third section of the piece (6:43—8:46). Rather than splicing textures according to onset events, the concluding section of the piece develops sound by altering the processing of voices, most easily heard from the use of tremolo effects applied to component sounds of chords (10:36—11:10).

Observations

In spite of availability of softwares and digital audio effects, the work on *Scenes* required original tools. Although the results of digital signal processing and acoustics researchers are of great use to the composer, creative work requires additional design considerations for utilization in music. *Trajectories* provided methods for exaggeration and attenuation of physically-modeled doppler shift and binaural effect. Exaggerated versions no longer represented a physical model, yet art and music often require unreal structures for aesthetic means.

The user interface for path manipulation was used to discover sonic behaviors. Rhythms and pitch effects emerging from aimless play—"babbling," as used by David Wessel—were necessary steps in learning how to design musical behaviors using sound spatialization.[29] Concealment of the numeric representations of the paths in favor of graphical metaphors of sound motion allowed the experiments to be driven by visual intuition and audition.

Trajectories was developed with binaural strategies suited for headphone experience of the work. The choice to write for headphones came from an observation that in spite of frequent experience of music in concert halls and live venues, headphone listening is the primary source of music experience—the privacy and immediacy of headphone listening are contributing factors. Although the presence of pinna filtering is lost during playback over speakers, it becomes invisible in such cases—the current stereo version of the piece is suited for stereo playback over pairs of loudspeakers and for live diffusion in multichannel environments. Experiments with live eight-channel diffusion of the piece and stereo audition over loudspeakers were carried out at the Center for New Music and Audio Technologies (CNMAT).

The scripts used to transform vocal sounds to glass textures provide a prototype mechanism for transforming rhythmic data into electroacoustic sound. Although the algorithmic approaches outlined in the work serve the materials of *Scenes* well, the tool would benefit from a dedicated user interface. Of particular desire are controls for manual insertion of textures to be modulated by the envelope, and additional envelopes for manipulating the textures' spectra computed from the spectral features of improvised sources. Expansion of *utterances* is planned for future work, to be informed by improvisatory approaches of other composers.

[29] David Wessel, "Designing Musical Instruments that Privilege Improvisation," filmed November 25, 2010, YouTube Video, 1:11:30, November 21, 2012, https://www.youtube.com/watch?v=uGASpqTXz4g. Discussion of babbling occurs at 10:50.

Conclusion

Composition of *Scenes* is a combination of imagined behaviors, embodied improvisation, and experimentation with tools. Vocal utterances mimicked digital processing. Digital arrangements were inspired by vocal articulations. Ideas discovered during the work on the last section found their way to inform the sounds of the introduction. Subtle mixing of noise to amplify the spatialization of high frequency sounds resulted in strategies for generating compelling timbres throughout the piece. The materials multiplied from musical play with a limited set of source sounds.

The tools created for the work apply simulated or vocal envelopes to glass sounds, generating rhythmic, articulated behavior. The work's obsession with the kinesthetic qualities of sound can be found throughout production: from the subtle animations of synthetic resonances, to spatialized sound, to vocal rhythms, to movement across sequenced textures, mixtures, beatings, and undulations. Technical work to realize the tools inspired manual construction of sounds by composing automation behaviors observed from the tools.

The value of the tools presented here, thus, is not of mere sound processing devices, readily available for a particular effect in the music of others. Rather, the tools are a manifestation of the act of music making, of composition, capable of transforming their author's ideas based on the lessons learned from conceiving and implementing them.

I would like to acknowledge Adrian Freed's contributions in lessons in human, aesthetic, and curious aspects of programming machines, Edmund Campion's *tempo-curving* work that inspired my own tools and softwares, and Myra Melford's unending support, help, and encouragement, particularly in enabling improvisation and performance. In addition, I would like to thank Rafael Valle for key contributions to onset detection work, and to Jonathan Kulpa for his guidance and patience. This work would not be the same without CNMAT, and without David Wessel's inspiring message to pursue construction of music with the rigor of scientific pursuit.

References

Ableton. "Learn More About Our Music Our Music Making Software Live | Ableton." https://www.ableton.com/en/live/ (accessed July 20, 2017).

Battier, Marc. “A Constructivist Approach to the Analysis of Electronic Music and Audio Art - between instruments and faktura.” Organized Sound 8(3). Cambridge University Press, 2003.

Battier, Marc. “What the GRM Brought to Music: From Musique Concrète to Acousmatic Music.” Organized Sound 12(3). Cambridge University Press, 2007.

Barrett, Natasha. “Spatio-musical Composition Strategies.” Organized Sound 7(3). Cambridge University Press, 2002.

Bello, Juan Pablo, Laurent Daudet, Samer Abdallah, Chris Duxbury, Mike Davies, and Mark B. Sandler. "A tutorial on onset detection in music signals." *IEEE Transactions on speech and audio processing* 13, no. 5 (2005): 1035-1047.

Bogdanov, Dmitry, Nicolas Wack, Emilia Gómez, Sankalp Gulati, Perfecto Herrera, Oscar Mayor, Gerard Roma, Justin Salamon, José R. Zapata, and Xavier Serra. "Essentia: An Audio Analysis Library for Music Information Retrieval." In *ISMIR*, pp. 493-498. 2013.

Ceccarelli, Luigi. Cadenza Esplosa. On Compendium International 2006 Bourges. Bourges: Mnémosyne Musique Media, 2006.

Continuum Analytics. "Welcome to Bokeh -- Bokeh 0.12.6 Documentation." http://bokeh.pydata.org/en/latest/ (Accessed July 27, 2017).

Cycling 74. "Max Software Tools for Media." https://cycling74.com/products/max (accessed July 20, 2017).

Dahlstedt, Palle. Gummi (Rubber) (1994-96). On E1999:1. Stockholm: Elektron Records, 2001.

Freed, Adrian. "Novel and Forgotten Current-steering Techniques for Resistive Multitouch, Duotouch, and Polytouch Position Sensing with Pressure." In *NIME*, pp. 230-235. 2009.

Freed, Adrian, and Andrew Schmeder. "Features and Future of Open Sound Control version 1.1 for NIME." In NIME, vol. 4, no. 06, p. 2009. 2009.

Gal, Andreas. "Node-wav" https://www.npmjs.com/package/node-wav (accessed July 20, 2017).

Gardner, Bill, and Keith Martin. "HRTF Measurements of a KEMAR Dummy-Head Microphone." http://sound.media.mit.edu/resources/KEMAR.html (accessed July 20, 2017).

GitHub. "Electron." https://electron.atom.io (accessed July 20, 2017).

Karlsson, Erik M. Un et Deux. On Compendium International 2006 Bourges. Bourges: Mnémosyne Musique Media, 2006.

Lehni, Jürg, and Jonathan Puckey. "Paper.js" http://paperjs.org (accessed July 20, 2017).

McClellan, Russel. "OSC-min: Simple Utilities for Open Sound Control in Node.js" https://www.npmjs.com/package/osc-min (accessed July 20, 2017).

Native Instruments. "Native Instruments - Software And Hardware For Music Production And Djing." https://www.native-instruments.com/en/ (accessed July 20, 2017).

NIME. "NIME | International Conference on New Interfaces for Musical Expression." http://www.nime.org (accessed July 20, 2017).

Normandeau, Robert. Murmures. On Robert Normandeau - Palimpsestes. Canada: Empreintes DIGITALes, IMED 12116, 2012.

Potisk, Tilen. "Head-Related Transfer Function." In Seminar Ia, Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana, Slovenia. 2015.

Pulkki, Ville. "Virtual sound source positioning using vector base amplitude panning." Journal of the audio engineering society 45, no. 6 (1997): 456-466.

Python Core Team. "Python: A dynamic, open source programming language." Python Software Foundation. https://www.python.org/ (accessed July 20, 2017).

Rostovtsev, Ilya Y. Source-bonding as a variable in electroacoustic composition: Faktura and acoustics in "Understatements". University of North Texas, 2010.

Schacher, Jan C., and Philippe Kocher. "Ambisonics spatialization tools for max/msp." Omni 500, no. 1 (2006).

Smalley, Denis. "Spectromorphology: explaining sound-shapes" Organized Sound 2(2). Cambridge University Press, 1997.

Waves. "Audio Plugins for Mixing, Mastering & Recording | Waves" http://www.waves.com (accessed July 20, 2017).

Wessel, David. “Designing Musical Instruments That Privilege Improvisation.” Filmed November 23, 2010. YouTube Video, 1:11:30. Posted November 20, 2012. https://www.youtube.com/watch?v=uGASpqTXz4g.

Wright, Matt. "The Open Sound Control 1.0 Specification." http://opensoundcontrol.org/spec-1_0 (accessed July 20, 2017).

Zölzer, Udo. Digital audio signal processing. John Wiley & Sons, 2008.

Appendix: Computing Spatial Movement

To compute the values of doppler shift, distance based attenuation, azimuth, and pan, source positions along the path are sampled at uniform intervals. Since the sound observer does not move, the doppler ratio is, simply:

$$r = \frac{c}{c + V}$$

where V is the velocity of the moving source (relative to the observer), c is the speed of sound, and r is the ratio of the observed frequency to the source frequency. The effect of this doppler shift on a sounding frequency f_0 is observed frequency $f = r \cdot f_0$

The change of amplitude, A, based on the distance from the listener, d, can be estimated by a square of the ratio of a chosen minimal distance, denoted d_{min}, at which the amplitude is not attenuated:

$$A = \left(\frac{d_{min}}{d}\right)^2$$

In the case of *trajectories*, the minimum distance is indicated by the distance value provided by the user. The radius of the safe-zone circle is used to translate the screen coordinates of the path to the world coordinates.

Azimuth, θ, is an angle value of the source to the listener, computed by the arctangent of the vector from the observer to the position of the source:

$$\theta = \arctan\left(\frac{y}{x}\right)$$

where x and y denote the components of the vector. Panning between the left and the right ears is a mapping of the azimuth angle, accounting for the traversal in front and in behind the observer.