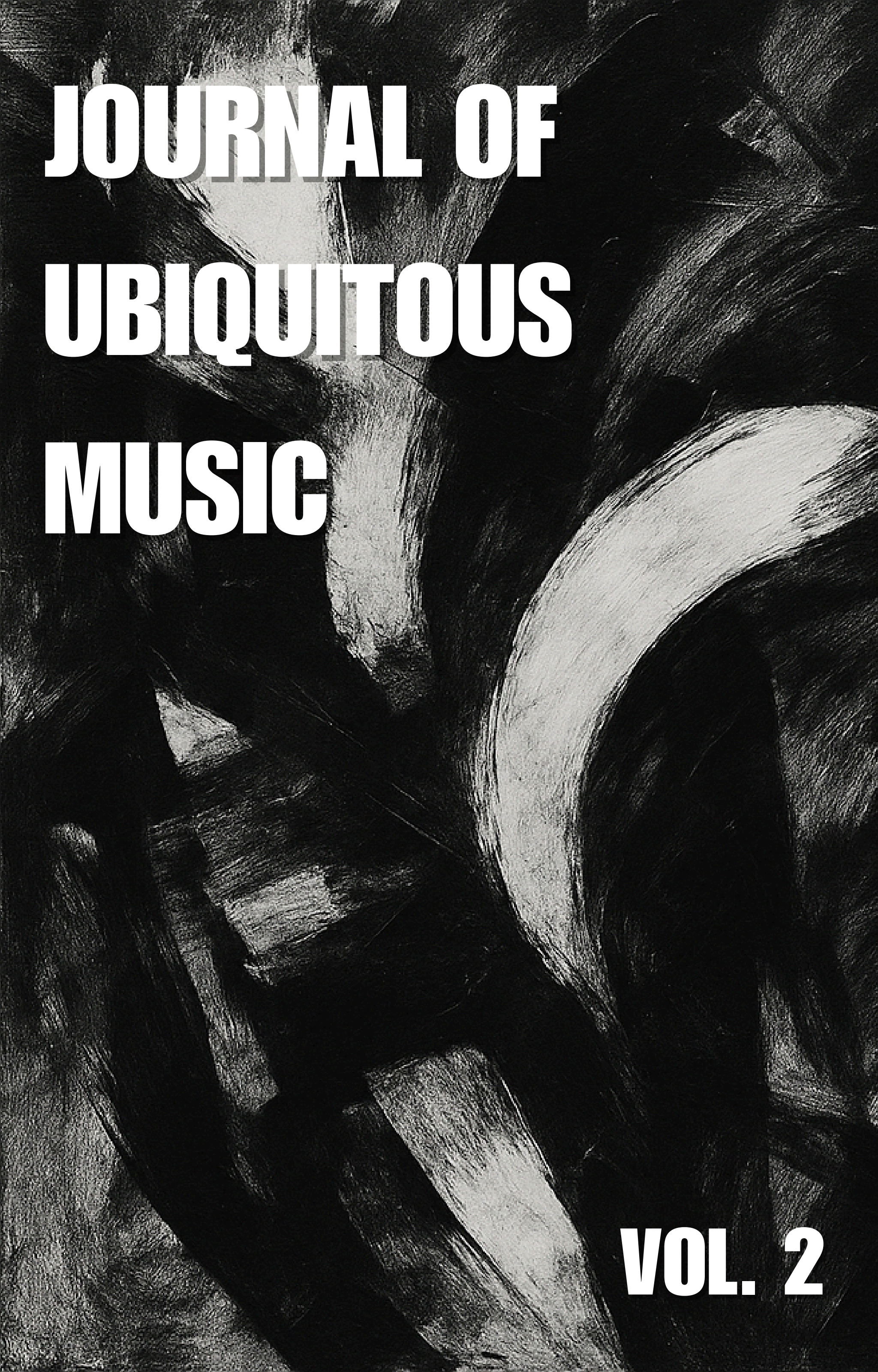


JOURNAL OF UBIQUITOUS MUSIC

VOL. 2

JOURNAL OF UBIQUITOUS MUSIC

JORNAL OF UBIQUITOUS MUSIC VOLUME 2 (2026)

DAMIÁN KELLER · ANDRÉ SONODA · LUZILEI[1] ALIEL

AMAZON CENTER FOR MUSIC RESEARCH
UBIQUITOUS MUSIC GROUP

EDITORIAL

Ubiquitous music research stands at the crossroads of multiple forces that have shaped the ways of thinking, designing, and deploying technological resources for post-2020 music-making. For us, the 2020 milestone is particularly significant because the period of the covid-19 pandemic highlighted the brittleness of the support infrastructure for musical interaction and the fragility of key 20th-century approaches to musical thought. The pandemic tsunami brought into focus the proposals laid out during the first wave of ubimus initiatives (2007-2014), suggesting that several of the emerging threads could be consolidated as sociotechnical frameworks. It is interesting to revisit some of the observations made by ubimus researchers in April 2020, when the lasting consequences of the pandemic period on artistic and educational practices were not as clear as today.

In their editorial Ubiquitous Music Making in COVID-19 Times, Keller, Costalonga and Messina (2020) formulate the following question: “will the new areas of ubimus application, highlighting the usage of domestic settings, the asynchronous strategies of group support and the incorporation of multiple modalities of exchange help to foster well-being, musical diversity and meaningful interaction? [...] For better or worse, [music-making] as we know it will have to make room for artistic practices compatible with a planet in decomposition [...]”. These aspects seem to have gained importance in post-2020 creative practice. Rather than the reactive attitude that permeated artistic initiatives of the last decades of the 20th century[1], ubimus practitioners tend to address environmental and social impact as central issues of technological design, which have to be considered before resources are deployed.

Another question that emerges from the confrontation with the extreme constraints triggered by the start of the pandemic is whether the extant infrastructure may provide the necessary support to bypass in-place music-making. “Despite the tendency to increase the amount of information available on the spot, overcoming some of the technical caveats faced by the telematic approaches to music-making such as network jitter and delay, there are human-performance limitations that even speed-of-light data transmission rates may not solve. One aspect is knowledge sharing. Network-based activities involving stakeholders with uneven levels of musical training demand careful attention to the strategies employed for supporting knowledge transfer [...]”[2]. This assertion is confirmed by several threads of post-2020 creative practice. A careless reader could interpret these words as subscribing to a network-only deployment of musical infrastructure. This is not the case. Web-based interaction has been supported since the early days of ubimus (see Miletto et al. 2011). But the design philosophy adopted by ubimus researchers has been expansive rather than resource-oriented[3]. Therefore, ubimus frameworks avoid platform-exclusive concepts that tend to get buried under the rubble of legacy technologies (typical examples are the Schaefferean ‘sound object’, the instrumental ‘score’ or the ‘laptop orchestra’, among several others)[4]. In this vein, it is interesting to note the recurring emergence of technofeudal myths. ‘Democratization’ is a case in point.

Among the objectives of ubimus research, reducing the digital divide seems to be at the top of the agenda. Nevertheless, providing access to technological resources does not necessarily change the dynamics of social exclusion. Consider, for instance, the incorporation of mobile platforms. It is true that mobile devices are currently available to significant segments of the population, even in low-income countries. This creates potential for creative usage, but it does not ensure an expansion of creativity. In particular, when the expanded availability of hardware is not accompanied by critical know-how, the effects of corporate infrastructure tend to be deleterious. As a result, the strategies applied in ubimus involve a policy of slow adoption: to avoid social disruptions, new resources are introduced after careful assessment of short-term effects and the consideration of long-term impact. This is akin to Weiser`s proposal of calm or slow technology (Weiser and Brown 1996), an approach to design that relegates attention-grabbing interaction techniques to specific demands and strives to incorporate peripheral-attention cognitive resources as material for design. Whole-body interaction is a case in point. Rather than approaching sound-shaping as an instrumentally oriented activity, ubimus frameworks explore body movements that do not require long periods of training or fine motor skills. The struck-string interaction framework discussed by Chakraborty et al. and by Kramann (in this volume) exemplifies this design approach.

Another thread highlighted by Keller, Costalonga, and Messina (2020) entails the use of improvisatory strategies in ubimus. “There is a well-established tradition of free-improvisatory practices in Brazil. But until recently, attempts to establish bridges between professionally oriented improvisation and the participation of lay musicians were rare. In an effort to overcome the artificial separation between musically trained subjects and casual collaborators, ubimus practitioners have laid out bridges to integrate improvisatory practices with active audience involvement [...]”. When approached from a ubimus-design perspective, the improvisational component is reconfigured: its functionality within the creative process is brought into question (Koszolko and Studley 2023). Given that ubimus is about ways of thinking, designing, and deploying, the incorporation of unplanned factors can be applied beyond sound-making. Contingencies are present in design (Aliel et al. 2024) and involve the incorporation of pliable musical concepts, such as musical stuff (Gómez Mejía et al. 2025). Therefore, improvisation is no longer understood just as synchronous sound-making with acoustic-instrumental resources. Contingency is a core component of thinking, designing, and deploying ubimus resources – hence it demands the reconfiguration of design methods. Interestingly, this aspect is explored in the hardware design techniques featured in Harding’s and Jagwani and

Lazzarini's proposals.

Another thread addressed by the authors of the 2020 editorial is ubimus dialogics. Based on the educational principles laid out by Paulo Freire (Shor and Freire 1987), dialogics highlights the role of horizontal exchanges among peers, fostering respect for cultural diversity and a positive attitude toward local knowledge. The participatory design movement in Scandinavia was strongly influenced by dialogics (Ehn 1988). These strategies involve an early incorporation of local stakeholders in the design process. This approach is exemplified in Santos, Defilippo, and Pimentel's project featured in this issue. Santos et al. focus on the demands of musical activities in the context of elementary education, with a special emphasis on the development of ubimus infrastructure applicable both to educational and domestic settings. They describe the implementation and usage of musical mats – an artifact composed of a desktop computer, a Makey Makey microcontroller, and a mat-based triggering mechanism. The prototype makes use of the shareware utility Soundplant2 as a tool for sonic production. Santos et al. carried out an informal study involving eighty seventh-grade students, focusing on classroom-based activities.

According to Keller et al. (2020) in the above-mentioned editorial, "the usage of interfaces and resources that emulate the behavior of European orchestral instruments is a prime example of genre-specific knowledge. Rather than calling this knowledge 'musical', it should be labeled 'orchestral' or even better, 'piano-', 'clarinet-', or 'guitar-based' knowledge. This view [of music-making] only targets the resources linked to the instrumental performance of acoustic and digitally emulated acoustic instruments. This type of knowledge has limited applicability since it does not encompass the rich experiences provided by a growing variety of multimodal artistic formats, by the application of analogue computing, and it does not engage with the recent contributions of the makers' movement to [music-making][...]". These three threads, multimodality, analogue computing, and support for DIY approaches to hardware design, are present in the current issue, highlighting a consistent theoretical path of ubimus developments. Furthermore, the observations regarding the specificity of timbre-related knowledge are now articulated as a set of ubimus frameworks, exemplified by struck-string interaction (see this issue's Section 1 with three contributions, by Chakraborty et al., Su et al. and Kramann).

Despite the efforts to avoid stylistic or genre-oriented biases, recent discussions in the ubimus community showcase a persistent concern to support legacy approaches to music-making. This is exemplified by frameworks such as the proposals of ubimus archaeologies and the musical-internet (IoMusT and IoMuSt)[5]. Lazzarini et al. (2023) document the recovery of one of the first music-programming compilers, MUSIC V (Mathews et al. 1969). A process of musicological investigation involving iterative adjustments of the programming environment enabled the deployment and synthesis, based on original code and materials, of Jean-Claude Risset's suite Little Boy (1969). Regarding the design for the musical internet, Fiorini et al. (2025) point to interesting contributions by Rich Gold, involving early implementations of networked resources. This thread may provide opportunities for further deployments of ubimus archaeologies, potentially unveiling musical experiences done at Xerox PARC in the 1990s.

In any case, acoustic-instrumental sonic resources present a difficult conundrum for legacy ubimus frameworks. Acoustic-instrumental sound is an important resource for music-making that cannot be excluded from ubimus designs. This body of material may impose some forms of sonic organization that are difficult to insert within genre-neutral approaches, such as equal-tempered tuning, functional tonality, pitch-based hierarchies, meter-based temporalities, or genre-oriented semantics. All of these are featured as central components of musical-information protocols or generative techniques and tend to reinforce the conservative tendencies of tool development. Thus, the incorporation of a genre-neutral technique to handle musical material is a relevant contribution to legacy frameworks. This initiative is exemplified by the application of spatialization processes on acoustic-instrumental sources. Peters, Koszolko, and Scott (this issue) report their work on the piece Immertio Overture. The novelty here lies in the hybrid approach to composing with instrumental sounds by incorporating sonic placement as a central dimension of creative thinking. The message is that the projected space of the sonic materials constitutes a core aspect of the creative process. How to tackle this dimension as an integrated parameter within the ubimus creative cycle is still an open question.

Summing up, the materials gathered in the second volume of the Journal of Ubiquitous Music feature original contributions that are well-aligned with the emergent threads of post-2020 creative practice. Cutting-edge technological innovation is exemplified by the incorporation of FPGA-based embedded computing that enhances the Ubimus Plugging Framework (Jagwani and Lazzarini, this issue; Keller, Jagwani, and Lazzarini 2025). Hardware prototyping is also exemplified by Santos et al.'s musical mats and Harding's analogue-synthesis modules (section 3). Multimodality is featured in the two Struck-String Interaction projects featured in section 1. An exploratory usage of spatialization is proposed by Peters et al. as a genre-neutral component of acoustic-instrumental sonic organizations. These proposals highlight a difficult coexistence between legacy and prospective frameworks, slowly emerging in ubimus post-2020 creative practice. A reconfiguration of musical first principles seems to be taking shape. Whether this will involve a dissolution of boundaries among inherited and forward-looking musical ways of thinking or a sharpened conflict between incompatible trends is an open issue to be resolved by careful and consistent application of ubimus methods. We look forward to the next chapters of this intriguing exploration.

ENDNOTES

[1] Undoubtedly soundscape composition provides a historical precedent regarding the need to consider not only the implications of musical activities within artistic venues, but also the educational potential of music-making released from the constraints of standard musical notation. The shortcomings of the soundscape approach as a tool for design are not necessarily conceptual, they are mostly methodological (see Lima et al. 2012; Gomes et al. 2014; Oleksik et al. 2008 for contrasting perspectives on this issue).

[2] The authors mention the widely adopted notion of ‘knowledge transfer’ but subsequent ubimus research indicates that nonintentional exchanges are as important as intentional exchange of musical information, hence the term ‘sharing’ seems to be better aligned to ubimus design.

[3] See the various criticisms of the notion of ‘thing’ in the series of publications dedicated to musical stuff (Messina, Keller, Aliel, Gómez Mejía, Filho and Simurra 2022; Messina, Stolfi, Aliel, Simurra and Keller 2024).

[4] There seems to be an acceleration of obsolescence in music-technological resources. This phenomenon may be due to the expansion of technofeudal hegemony in the central countries. But it may also be caused by an increased fragility of the ways of doing in the peripheral communities of practice. A brittle material base tends to disrupt knowledge-sharing strategies. Thus, there may be a synergy in the obsolescence processes: lack of material support tends to constrain knowledge-sharing and lack of know-how compromises the adoption of infrastructure.

REFERENCES

Aliel, L., Simurra, I., Messina, M., & Keller, D. (2023). Ecomprovisation: Project Markarian 335. Organised Sound, 28(3), 328–337. https://doi.org/10.1017/S1355771823000651

Ehn, P. (1988). Work-oriented design of computer artifacts. Stockholm, Sweden: Arbetslivscentrum.

Gomes, J., Pinho, N., Lopes, F., Costa, G., Dias, R., Tudela, D., & Barbosa, Á. (2014). Capture and transformation of urban soundscape data for artistic creation. Journal of Science and Technology of the Arts, 6(1), 97–109. http://artes.ucp.pt/citarj/article/view/132

Gómez Mejía, C. M., Keller, D., Alliel, L., & Messina, M. (2025). Ubimus, post antropocentrismo y la ubicua inmaterialidad del cachivache musical. Cuadernos de Música Iberoamericana, 38, 49–72. https://doi.org/10.5209/cmib.98867

Keller, D., Costalonga, L., & Messina, M. (2020). Editorial: Ubiquitous Music Making in COVID-19 Times. Proceedings of the Workshop on Ubiquitous Music (UbiMus 2020), 3–16. https://doi.org/10.5281/zenodo.4245850

Keller, D., Jagwani, A., & Lazzarini, V. (2025). The Ubimus Plugging Framework: Deploying FPGA-Based Prototypes for Ubiquitous Music Hardware Design. Computers, 14(4), 155. https://doi.org/10.3390/computers14040155

Koszolko, M. K., & Studley, T. (2023). From Site-Specific Sampling to Gamification: An exploration of performative engagement with the environment. Organised Sound, 28(3), 338–351. https://doi.org/10.1017/S1355771823000547

Lazzarini, V., Keller, D., & Radivojević, N. (2023). Issues of ubiquitous music archaeology: Shared knowledge, simulation, terseness, and ambiguity in early computer music. Frontiers in Signal Processing, 3, 1132672. https://doi.org/10.3389/frsip.2023.1132672

Lima, M. H., Keller, D., Pimenta, M. S., Lazzarini, V., & Miletto, E. M. (2012). Creativity-centred design for ubiquitous musical activities: Two case studies. Journal of Music Technology and Education, 5(2), 195–222. https://doi.org/10.1386/jmte.5.2.195_1

Mathews, M. V. (1969). The Technology of Computer Music. Cambridge, Massachusetts: MIT Press.

Messina, M., Keller, D., Aliel, L., Mejía, C. M. G., Filho, M. C., & Simurra, I. (2022). The Internet of Musical Stuff (IoMuSt): Ubimus perspectives on artificial scarcity, virtual communities and (de)objectification. INSAM: Journal of Contemporary Music, Art and Technology, (9), 26–50. https://doi.org/10.51191/issn.2637-1898.2022.5.9.26

Messina, M., Stolfi, A. de S., Aliel, L., Simurra, I., & Keller, D. (2024). The Internet of Musical Stuff: Towards an Aesthetically Pliable Musical Internet. International Journal of Software Innovation, 12(1), 1–19. https://doi.org/10.4018/ijsi.344018

Miletto, E. M., Pimenta, M. S., Bouchet, F., Sansonnet, J.-P., & Keller, D. (2011). Principles for music creation by novices in networked music environments. Journal of New Music Research, 40(3), 205–216. https://doi.org/10.1080/09298215.2011.603832

Oleksik, G., Frohlich, D., Brown, L. M., & Sellen, A. (2008). Sonic interventions: Understanding and extending the domestic soundscape. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’08, 1419–1428. https://doi.org/10.1145/1357054.1357277

Shor, I., & Freire, P. (1987). What is the “dialogical method” of teaching? Journal of Education, 169(3), 11–31.

Weiser, M., & Brown, J. S. (1996). Designing calm technology. Powergrid Journal, 1(1).

SECTION I - ADVANCES IN STRUCK-STRING INTERACTION

Vitruvian Whole-Body Creative Action: A Proposal for the Struck-String Interaction Framework (Chakraborty; Keller; Timoney);

Mandala Music – Description Installation Workshop (Kramann);

MIDI Adaptive Tuning Strategies by Means of AI-Based Struck-String Interaction in Ubimus (Su; Timoney; Keller).

VITRUVIAN WHOLE-BODY CREATIVE ACTION:

A PROPOSAL FOR THE STRUCK-STRING INTERACTION FRAMEWORK

SUTIRTHA CHAKRABORTY · DAMIÁN KELLER · JOSEPH TIMONEY[1]

ABSTRACT

The Vitruvian Creative-Action Metaphor is proposed as an approach to tackle whole-body interaction with piano-like sounds. We report on the design of a visual-tracking adaptive mechanism, implemented as a camera-based system deployable on browser technology. Our presentation documents the initial prototypes, available for usage by the ubimus community.

1 INTRODUCTION

Recent advances in ubimus research point to the emergence of an initiative targeting the development of infrastructure and the exploration of piano timbre without resorting to the acoustic-instrumental concepts tied to interacting with keyboards.

This tendency to equate musical interaction to the usage of acoustic instruments has not diminished, despite the significant changes implied by the adoption of mobile, embedded and network-based infrastructure. Given this state of affairs, various initiatives within our community are starting to

---

[1] Ubiquitous Music Group - October 2024 - March 2025

unveil concepts and techniques that offer an alternative to the acoustic-instrumental perspective to deal with piano sounds: the struck-string interaction framework [Kramann and Keller, 2024], [Su et al., 2024].

Struck-string interaction involves a loose set of strategies for handling piano sounds, with a strong emphasis on the reduction of the cognitive load and the temporal investment typically demanded by the virtuosic mindset of the instrumental approach. We are interested in expanding the musical possibilities of the parametric control of piano-like sonic resources, encompassing both in-place and remote opportunities for shared musical experiences. Hence, our proposal involves revisiting well-established practices and also fostering an overhaul of what is understood as ”piano music”.

In this paper, we introduce the Vitruvian Creative-Action Metaphor within the context of struck-string interaction (henceforth ssi). The remaining sections encompass a presentation of the key characteristics of the Vitruvian Metaphor, a documentation of the computational components, highlighting the potentials and caveats of the current imple- mentation. We provide details of the system architecture, the adaptive components and their interplay.

2 RELATED WORK

Gesture- and body-driven mappings to musical sound have been investigated across various sensing modalities. Markerless, camera-based interfaces paired with interactive machine learning frameworks have enabled performers to map full-body motion to audio and visuals. For example, Schedel et al. demonstrated using the Wekinator toolkit to translate Microsoft Kinect skeletal data into musical and visual outputs for the ensemble 000000Swan [Schedel et al., 2011]. Wearable IMU-based gloves, typified by the Mi. Mu Gloves designed by Imogen Heap's team, employ flex and orientation sensors to capture fine-grained hand and arm movements for expressive, low-latency control[2]. Non-contact ultrasonic systems have also been used for gesture recognition: Sang et al. proposed an ultrasonic active sensing array for hand-gesture classification, achieving a performance benchmark suitable for interactive applications [Sang et al., 2017].

These systems feature a mapping layer that transforms raw sensor data into sound synthesis or parametric control. Hunt and Wanderley introduced a two-layer model (a) separating sensor-specific feature extraction from synthesis parameter assignment (b) to decouple interface design from sound generation [Hunt and Wanderley, 2002]. Building on this, Verfaille et al. formulated a multi-level mapping framework for au- dio processing, distinguishing gestural control from adaptive, sound-driven modulation [Verfaille et al., 2006]. Our Struck-String Interaction (ssi) framework extends these approaches by adopting a Vitruvian Man–inspired creative-action metaphor, unifying discrete (inner-zone) and continuous (outer-zone) mappings in a browser-native, camera-based computational environment.

3 THE VITRUVIAN CONCEPT

Our starting point consists of stripping away the

---

[2] https://mimugloves.com

requirements of piano-sound deployments to a bare minimum. This Ockham-Razor exercise prompts us to tackle interaction support as a three-component ecosystem: a human body that generates information by means of movements, a tracking component that translates these movements into data, and a rendering component that furnishes the sonic outcomes and informational feedback necessary to adjust the participants' behaviours to local demands and aesthetic goals. We emphasise that no genre prescriptions are adopted other than a focus on piano-related sounds (cf. [Chakraborty et al., 2022] and [Keller et al., 2023] for a similar approach ap- plied in banging interaction).

A further constraint of our proposed design entails enhancing the sustainabil- ity of infrastructure. As proposed by other ubimus frameworks, we employ browser- based technologies that are compatible with stationary, embedded or mobile devices [Lazzarini et al., 2014] [Lazzarini et al., 2020] [Yi and Letz, 2020].

Drawing inspiration from Da Vinci's (1487) painting, The Vitruvian Man - which epitomizes human-body proportions through geometric ratios - the proposed creative- action metaphor defines spatial interaction zones to determine the type and character of musical outcomes. The Vitruvian Metaphor allows participants to perform struck-string sonic models by means of simple body movements. It employs a webcam to enable MediaPipe-based pose detection and it incorporates MIDI protocols for parametric control and audio synthesis. Key features include the following items.

- **Pose Tracking:** Synchronous detection of body, hands, and face landmarks.

- **Movement Tracking and Gesture Detection:** Algorithms to compute motion-speed, direction and positional changes.

- **Vitruvian Visuals:** A dynamically scaled square with inner and outer circular zones aimed at whole-body interaction.

- **MIDI Integration:** Configuration of MIDI channels and parametric control of piano-like audio-synthesis models.

- **Visual Feedback:** On-screen graphics that display geometric containers, labels, and colour-coded visual anchors to complement proprioceptive cues.

4 SYSTEM ARCHITECTURE

The current implementation is structured as a real-time processing pipeline that converts video input into musical output, furnishing visual feedback of body actions. The primary components are as follows.

4.1 OVERVIEW

Figure 1 depicts the high-level data flow.

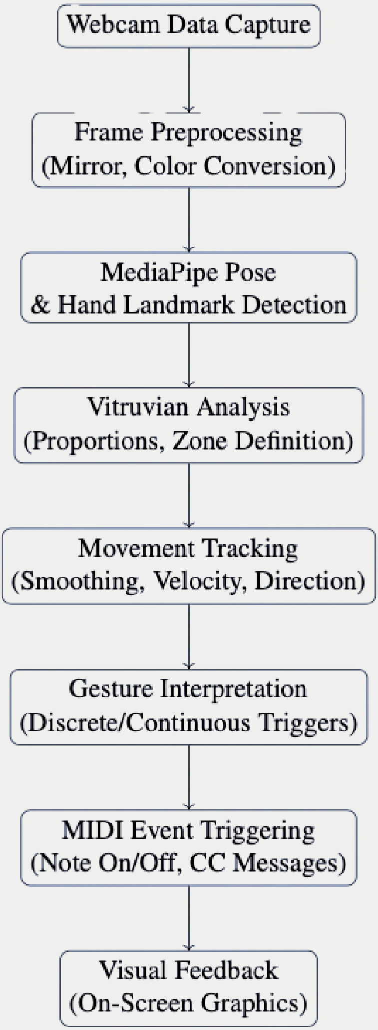


FIGURE 1: VITRUVIAN PROTOTYPE WORKFLOW: FROM VIDEO CAPTURE TO MIDI OUTPUT AND VISUAL FEEDBACK

4.2 DETAILED COMPONENTS

- **Webcam Capture:** Utilizes OpenCV to obtain a live video feed.

- **Frame Preprocessing:** Applies mirroring and colour space conversions.

- **Pose Detection:** Leverages MediaPipe's holistic model to extract 33 body tokens, 21 per hand tokens, and 468 face tokens.

- **Vitruvian Analysis:** Computes a dynamic Vitruvian container based on the user's height (distance from head to feet) and overlays a square with two concentric circles defining interaction zones.

- **Movement Tracking:** Uses a MovementTracker class (employing a circular buffer via Python's deque) to calculate speed, direction, and trigger conditions.

- **Gesture Interpretation:** Differentiates between continuous (e.g., sustained index finger position) and discrete (e.g., foot or wrist movement) gestures.

- **MIDI Communication:** Sends MIDI messages through a virtual port using the mido library, mapping gestures to different channels.

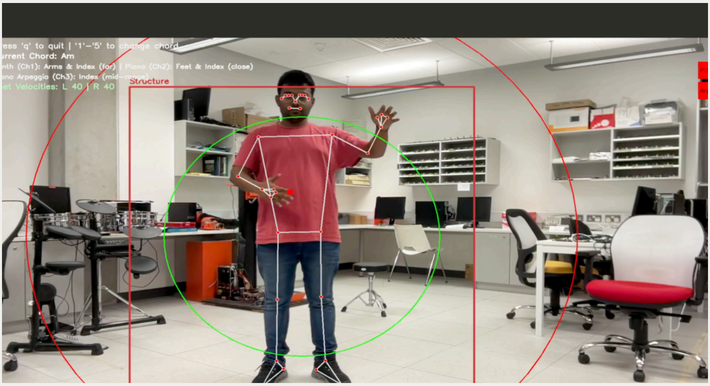
- **Visual Feedback:** Renders real-time graphics – including the Vitruvian container, zones, labels, and colour panels – through the video feed.

5 POSE DETECTION WITH MEDIAPIPE

Pose detection is a core functionality. MediaPipe's Holistic model is configured with a minimum detection confidence of 0.6 and a tracking confidence of 0.6 to robustly extract various landmarks. A key feature of the model is to keep a consistent tracking behaviour, regardless of the distance between the device and the subject. Thus, within the camera's field of vision, whole-body motion is supported.

- **Body Landmarks:** Nose, shoulders, hips, wrists, and feet determine the overall pose and Vitruvian scaling.
- **Hand Landmarks:** The right-hand index fingertip is crucial for fine gesture con- trol.
- **Face Landmarks:** Although not directly used for MIDI mapping, they enhance tracking reliability.

Figure 2 (a placeholder diagram) illustrates the key landmark groups.



```
ess 'q' to quit | '1'–'5' to change chord
urrent Chord: Am
nth (Ch1): Arms & Index (far) | Piano (Ch2): Feet & Index (close)
ano Arpeggio (Ch3): Index (mid–range)
et Velocities: L 40 | R 40
```

FIGURE 2: SYNCHRONOUS MOTION CAPTURE AND INTERACTION INTERFACE FROM THE VITRUVIAN CREATIVE-ACTION METAPHOR. THE PERFORMER IS TRACKED USING SKELETAL KEYPOINTS, WITH DISTINCT ZONES (HIGHLIGHTED IN RED AND GREEN) MAPPED TO VARIOUS PIANO-LIKE SOUND MODELS. THIS SYSTEM ENABLES EXPRESSIVE FULL-BODY MUSICAL INTERACTION THROUGH SPATIAL BODY-MOTION RECOGNITION

6 MOVEMENT TRACKING AND GESTURE DETECTION

The system relies on accurately tracking movements and interpreting gestures. A dedicated *MovementTracker* class processes sequential landmark data to compute speed of motion and direction.

6.1 FEET AND WRIST TRACKING

Feet: Movements are tracked using a 15 position buffer.The average y-coordinate difference (speed of movement) is mapped to MIDI velocity. Direction changes (e.g., upward vs. downward movement) trigger piano onset events on MIDI Channel 2.

6.2 INDEX FINGER TRACKING

The right-hand index finger provides nuanced control based on both its position and movement:

- Crossing Detection: A function computes the side of a point relative to a line (from the right shoulder to the right hip) using the following cross product.

$$side=(Bx-Ax)(Py-Ay)-(By-Ay)(Px-Ax)$$

A sign change indicates that the index finger has crossed a defined boundary.

- **Distance Zones:** The distance from the body centre is used to define three zones.

***Far Zone* (outside of the outer circle):** Activates a continuous synth-pad note (Channel 1).

***Close Zone* (inside the inner circle):** Triggers a piano event (Channel 2).

***Intermediate Zone* (between the circles):** Initiates a piano arpeggio (Channel 3).

Figure 3 outlines the movement-tracking algorithm.

7 VITRUVIAN VISUALS AND PROPORTIONAL ANALYSIS

Inspired by Da Vinci's Vitruvian Man, the system dynamically computes a Vitruvian Size based on the distance between the subject's head and feet. This measurement scales a square - the Vitruvian container – drawn on the screen. Within this square, two concentric circles are overlaid.

- **Inner Circle:** Defines a zone for discrete, percussive triggers.
- **Outer Circle:** Establishes the boundary for continuous sound control.

Figure 4 illustrates the geometric layout.

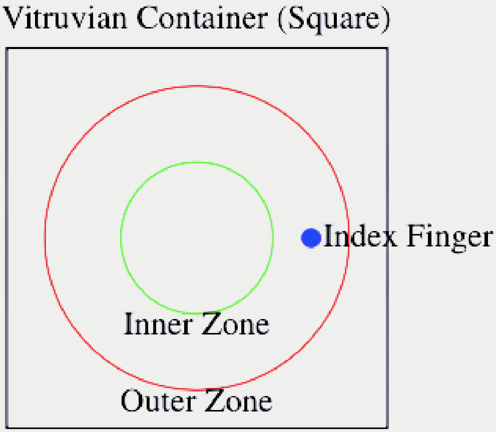


FIGURE 3: VITRUVIAN VISUALS: A SQUARE AND INNER AND OUTER CIRCULAR ZONES DEFINING GESTURE-BASED CONTROL AREAS

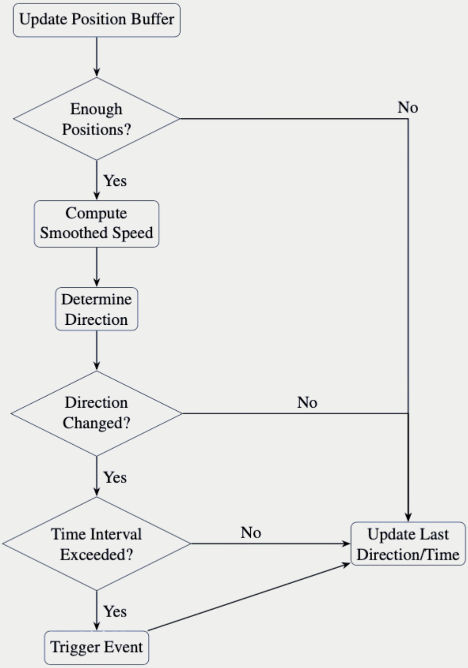


FIGURE 4: MOVEMENT TRACKER ALGORITHM: PROCESSING POSITION DATA TO TRIGGER MIDI EVENTS

8 MIDI COMMUNICATION AND SONIC RENDERING

The system converts body gestures into musical output by means of MIDI events. MIDI messages are implemented using the Python mido library, routed over a virtual MIDI port.

8.1 MIDI CHANNEL MAPPING

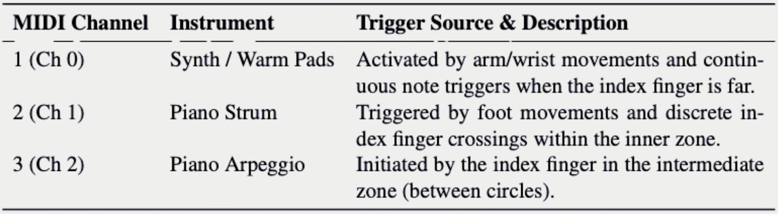
MIDI channels are assigned to various sound types:

- **Channel 1 (mido channel 0):** Synth or warm pad sounds. Triggered by arm/wrist movements and continuous note mode (when the index finger is far).

- **Channel 2 (mido channel 1):** Piano events. Activated by feet movements or when the index finger crosses into the inner zone.

- **Channel 3 (mido channel 2):** Piano arpeggios. Initiated when the index finger is positioned between the inner and outer circles.

Table 1 summarizes the parametric MIDI-based mapping.



MIDI Channel	Instrument	Trigger Source & Description
1 (Ch 0)	Synth / Warm Pads	Activated by arm/wrist movements and continuous note triggers when the index finger is far.
2 (Ch 1)	Piano Strum	Triggered by foot movements and discrete index finger crossings within the inner zone.
3 (Ch 2)	Piano Arpeggio	Initiated by the index finger in the intermediate zone (between circles).

TABLE 1: MIDI CHANNEL MAPPING AND ACTIONS

8.2 MIDI WORKFLOW DIAGRAM

The following diagram (Figure 5) illustrates the mapping from gesture detection to MIDI event dispatch.

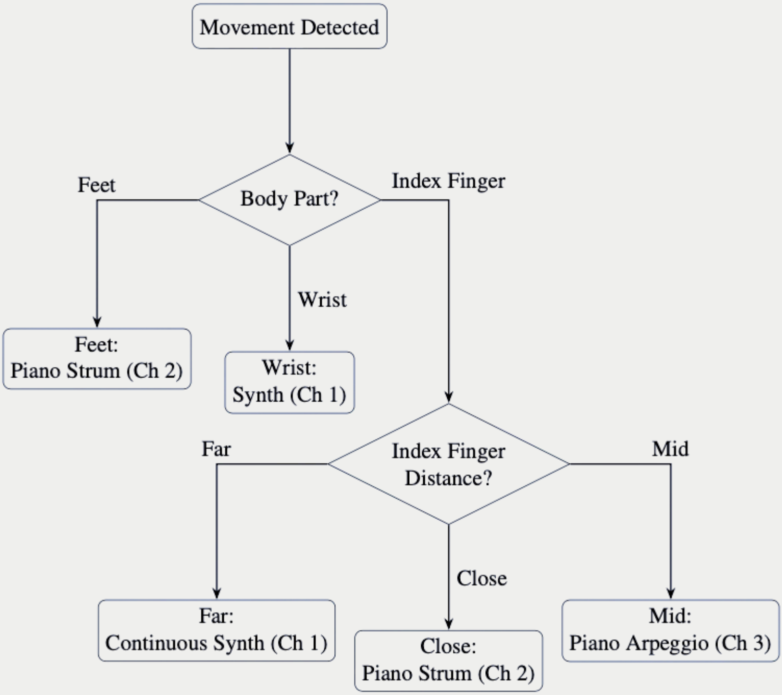


FIGURE 5: MIDI TRIGGER MAPPING: FROM BODY PART DETECTION TO SOUND GENERATION ON DESIGNATED CHANNELS

8.3 MOVEMENT–TO–SOUND MAPPING RATIONALE

Each zone's behaviour of prototype v.1 draws on common practice in gesture-based de- sign.

- **Inner Zone (0–30% radial range):** discrete index-finger crossings trigger piano onsets, leveraging the fine motor affordances of finger taps.

- **Intermediate Zone (30% - 60%):** index-finger position modulates arpeggios, preserving pitch continuity while allowing control of melodic contour.

- **OuterZone (60% - 100%):** sustained arm/wrist motion controls synth-pad timbre and volume, aligning continuous spatial gestures with sustained events.

We frame these choices as hypotheses grounded in gesture-based common practice, to be validated in future empirical studies.

8.4 MIDI MODULE

The MIDI module tackles the basic functionality for the generation of piano-like events.

1. **Initialization:** A virtual MIDI port is opened (e.g., using mido.open output).

2. **Program Changes:** Instrument settings are configured via program change messages (e.g., Program 90 for synths, Program 1 for piano).

3. **Onset Triggering:** On detecting a gesture, a note on message (with velocity proportional to movement speed) is sent, often accompanied by control change (CC) messages (e.g., reverb on CC 91 or modulation on CC 1).

4. **Note Off:** After a predetermined duration for discrete events, a note off mes- sage is dispatched.

9 VISUAL FEEDBACK

Visual feedback reinforces the interaction by overlaying dynamic graphics onto the video feed.

- **Vitruvian Container:** A square scaled to the user's height, representing ideal proportions.
- **Concentric Zones:** Two circles (inner and outer) demarcate regions for discrete and continuous triggers.
- **Labels and Colour Panels:** On-screen text displays pitch-aggregates names, movement velocities, and MIDI pitches, while a colour panel maps MIDI octaves to specific colours (e.g., blue for lower octaves, red for higher octaves).

10 VITRUVIAN CONCEPTS AND THEIR IMPLEMENTATION

The system's aesthetic and functional design is inspired by Leonardo da Vinci's The Vitruvian Man.

- **Proportional Scaling:** The distance from the head to the feet is used to dynamically size the Vitruvian container (square), ensuring the visual feedback remains proportional to the user.
- **Zone Allocation:** Two concentric circles within the square define distinct interaction areas:

  Inner Zone: Reserved for discrete, percussive triggers (e.g., piano strums or onsets).

Outer Zone: Engages continuous sound control (e.g., sustained synth-pad sounds).

- **Symbolism:** The geometric forms not only serve a functional role but also suggest embodiment of musical actions, highlighting their relationship to human anatomy.

Figure 6 illustrates the integration of the visual Vitruvian components.

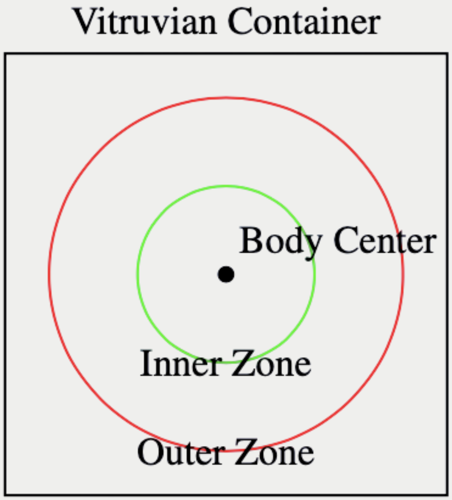


FIGURE 6: VITRUVIAN FRAMEWORK: DYNAMIC SCALING AND ZONE ALLOCATION INSPIRED BY MODELLED HUMAN PROPORTIONS

10.1 CAMERA-BASED TRACKING: RATIONALE AND LIMITATIONS

We chose a browser-native, webcam-only approach for maximum accessibility. These are some of the trade-offs.

- **Latency:** Typical per-frame delay of 30-50 ms at 30 fps (mitigated by downsam- pling to 15 fps for tracking loops).

- **Orientation Constraint:** Performers must face the camera; we mitigate this via on-screen alignment guides and wide-angle lens correction.

- **No Wearables:** Unlike IMU gloves, our system requires zero setup after opening a web-based interface.

11 ALGORITHMS AND IMPLEMENTATION DETAILS

Several key algorithms enable the conversion of body movements into musical outputs.

11.1 MOVEMENTTRACKER ALGORITHM

The *MovementTracker* class employs a circular buffer (using *Python's deque*) to store recent positions and computes:

- **Speed:** Calculated as the average Euclidean distance between successive positions:

$$\text{speed} = \frac{1}{n-1} \sum_{i=1}^{n-1} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}$$

- **Direction:** Inferred by comparing the first and last positions in the buffer.

- **Trigger Conditions:** A note is triggered only if a significant change in direction occurs and a minimal time interval (e.g., 0.25 seconds) has passed.

11.2 DISTANCE AND CROSSING CALCULATIONS

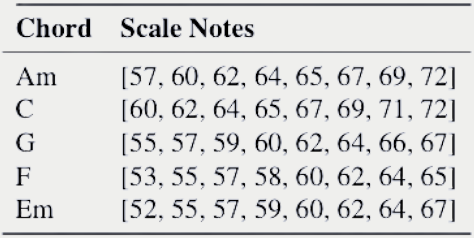
- **Distance from Center:** The Euclidean distance of the index finger from the body center is computed as:

$$\text{dist} = \sqrt{(x_f - x_c)^2 + (y_f - y_c)^2}$$

- **Point-Line Crossing:** Using the cross product method (as shown earlier), the system determines when the index finger crosses a defined boundary.

11.3 PITCH SELECTION AND SCALE MAPPING

Musical events are selected based on user-defined scales and chords. Table 2 lists example scales employed by the current prototype.



Chord	Scale Notes
Am	[57, 60, 62, 64, 65, 67, 69, 72]
C	[60, 62, 64, 65, 67, 69, 71, 72]
G	[55, 57, 59, 60, 62, 64, 66, 67]
F	[53, 55, 57, 58, 60, 62, 64, 65]
Em	[52, 55, 57, 59, 60, 62, 64, 67]

TABLE 2: EXAMPLE MUSICAL SCALES FOR PITCH SELECTIONS

12 PRELIMINARY DISCUSSIONS: CAVEATS AND POSSIBILITIES

Musical events are selected based on user-defined scales and chords. Table 2 lists example scales employed by the current prototype. We report on the design of the Vitruvian Creative-Action Metaphor - proposing a prototype based on computer vision tracking, synchronous pose tracking and MIDI connectivity to turn human gestures into whole-body musical experiences. Inspired by Leonardo da Vinci's The Vitruvian Man, human-body proportions and movements are captured through adaptive techniques based on a combination of MediaPipe's holistic landmark detection, custom movement-tracking and a MIDI-based communication.

In this report we described the prototype's architecture, algorithmic approaches and the ubimus strategies required to handle piano-like sounds. We have shared our tool within a research focus-group. A working prototype can be accessed online. Our aim with this presentation is to gather feedback from the ubimus community on the proposed concepts, the design choices and the targeted types of musical experiences.

During our discussions toward the early deployments of the Vitruvian prototype, two approaches to implementation emerged: visual tracking and accelerometer-based sensing. Both techniques feature advantages and limitations within the context of struck-string interaction. Arguably, visual tracking might present some caveats when dealing with finger movements. This is an area of active research and our demonstrations will feature problems and solutions.

Another aspect to consider is the increased cognitive demands of multiple simultaneous body movements while tackling complex musical events, such as pitch-aggregates that feature multiple dynamic levels in sequence-based interactions. We address strategies to tackle these issues.

The Vitruvian Metaphor is proposed as an approach to struck-string interaction that furnishes a simplification of means for parametric control which is deployable on browser-based technology. The implemented prototype supports interaction with piano- like sounds through visual tracking of body movements. Alternative approaches can complement our proposal. We look forward to suggestions and insights from the ubimus community to expand and refine the vitruvian design.

13 FUTURE WORK

Future work involves quantifying the cognitive load of multi-limb gestures using assess- ment tools such as the NASA-TLX and dual-task techniques. We will also tackle artistic applications, collecting performance recordings, participants feedback, and expert analyses, to gauge its creative caveats and potentials. By profiling the required skill levels, we will develop adaptive gesture-to-sound techniques tailored for novices and experts. We will also incorporate frameworks from ubimus theory to organize and justify our movement→sound assignments and interaction-design decisions. Finally, to support multiple approaches to temporalities, we will optimize the tracking algorithms and adjust the WebRTC settings to drive end-to-end latency below 20 ms.

REFERENCES

Chakraborty, S., Yaseen, A., Timoney, J., Lazzarini, V., and Keller, D. (2022). Adaptive touchless whole-body interaction for casual ubiquitous musical activities. In Proceedings of the International Computer Music Conference (ICMC2022), pages 132–138. Limerick, Ireland: University of Limerick.

Hunt, A. and Wanderley, M. M. (2002). The importance of parameter mapping in electronic instrument design. Organised Sound, 7(2):97–108.

Kramann, G. and Keller, D. (2024). Struck-string interaction: Exploring piano timbre beyond the piano. In Proceedings of the Ubiquitous Music Symposium (UbiMus 2024). Macau, China: Ubiquitous Music Group.

Lazzarini, V., Costello, E., Yi, S., and ffitch, J. (2014). Development tools for ubiquitous music on the world wide web. In Ubiquitous Music, Computational Music Science, pages 111–128. Heidelberg Berlin: Springer International Publishing.

Lazzarini, V., Keller, D., Otero, N., and Turchet, L. (2020). Ubiquitous Music Ecologies. London: Taylor & Francis (Routledge).

Sang, Y., Shi, L., and Liu, Y. (2017). Micro hand gesture recognition system using ultrasonic active sensing. arXiv preprint arXiv:1712.00216.

Schedel, M., Perry, P., and Fiebrink, R. (2011). Wekinating 000000swan: Using machine learning to create and control complex artistic systems. In Proceedings of the International Conference on New Interfaces for Musical Expression (NIME), pages 453–456, Oslo, Norway.

Su, R., Timoney, J., and Keller, D. (2024). Just intonation and inharmonic- ity in struck-string interaction. In Proceedings of the Ubiquitous Music Symposium (UbiMus 2024). Macau, China: Ubiquitous Music Group.

Verfaille, V., Wanderley, M. M., and Depalle, P. (2006). Mapping strategies for gestural and adaptive control of digital audio effects. Journal of New Music Research, 35(1):71–93.

Verfaille, V., Wanderley, M. M., and Depalle, P. (2006). Mapping strategies for gestural and adaptive control of digital audio effects. Journal of New Music Research, 35(1):71–93.

Yi, S. and Letz, S. (2020). The browser as a platform for ubiquitous music. In Lazzarini, V., Keller, D., Otero, N., and Turchet, L., editors, Ubiquitous Music Ecologies, pages 185–206. London: Routledge.

MANDALA MUSIC – DESCRIPTION INSTALLATION WORKSHOP

GUIDO KRAMANN[1]

KRAMANN@TH-BRANDENBURG.DE

ABSTRACT

Mandala Music is a composition method for amateurs. Similar to sudoku, but in a larger area and according to rules that allow more freedom, the user enters numbers from 0 to 9 in a square matrix. Through continuous updates, the resulting number structure is transformed into a musical composition and provided to the user as feedback. The rules of Mandala Music require the user to create rhythmic patterns with the numbers, which should also be reused in parts if possible. In this respect, the rules force the user into musical thought patterns and challenge him to act creatively within these thought patterns.

1 INTRODUCTION

There have been and still are many approaches to transforming the course of a game into music, including very prominent ones such as John Cage’s sonification of chess games as a happening [Cross 1999]. In general, these examples do not claim to be tools for comprovization, or even for composing. Typically, the game and its rules exist beforehand and an automatic musical accompaniment to the game is implemented, the style of which at best adapts to the

---

[1] Technische Hochschule Brandenburg.

game situation, see e.g. [Hamilton 2014].

On the other hand, there are already many approaches to converting non-musical, voluntary movements into music, for example on the basis of sensory physical gestures, see e.g. [Zbyszynski et al. 2021]. In contrast to the approach with sonified gameplay, the resulting music is not merely a by-product of the independent gameplay, but the user can learn via the musical feedback to use his gestures to make a certain musical statement, and via embodiment (see also [Merleau-Ponty and Schroeder 1996]) this approach is even particularly predestined to be able to realize particularly emotional and musically rich performances. This is a useful and widely used approach for comprovization. It is less useful for composing, as the musical event is ephemeral and difficult to reproduce. This in itself would not be sufficient reason to reject the approach for composing. However, this approach denies the user an essential mode of action for composing: Due to the lack of reproducibility and the inability to record anything, it is not possible to make incremental, small, evolutionary improvements to a musical structure in the making.

1.1 A THIRD WAY

With Mandala Music, a third way is proposed here compared to the musicalization of games and the musicalization of physical gestures, in which the focus corresponds more to the requirements of composing and less to those of performance or improvisation, and which aims to make this skill accessible to laypeople in an intuitive way.

(Gaming) rules are more than just setting boundaries. In fact, they are the reason to become active in a game in the first place. They challenge the user to prove that their intellect is capable of mastering the game while adhering to the rules. This fact is exploited in Mandala Music on the one hand to avoid the horror vacui, i.e. the helplessness that sets in when you look at a blank musical score. On the other hand, the starting point for the concept of Mandala Music was the requirement to implement compositional rules directly as rules in a game in order to allow game decisions to become immediate compositional decisions by the user. When implemented as a computer program, it is possible to automatically monitor whether the rules are being adhered to. The rules of the game take on the role that music theory plays in composition. The course of the game is then the production of the composition on the basis of a specific music theory.

1.2 CLASSIFICATION OF MANDALA MUSIC IN THE MODERN COMPOSITION LANDSCAPE

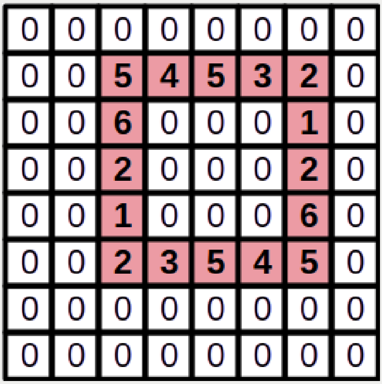
First of all, it is important to clarify what is actually meant by “composing a piece of music” in the context of Mandala Music. A modern composer of New Music would at most use the system presented here to get inspiration for the musical structure to be created, but would certainly be far from calling what is created in Mandala Music a “composition”. This is because in New Music, valid conventions are typically called into question with every new composition. Modern composers develop their own tonal language. Conventions and rules retain their validity at most over the course of a limited cycle of successive works (Sofia Gubaidulina, Karlheinz Stockhausen, Giörgy Ligeti), or are at least tied to the person of the composer (Arvo Pärt, Philip Glass), see [Gottstein 2024], pp. 92-98, 114-124, 151-155, 227-232. In the case of younger composers, one would not want to make such an assessment at an early stage. Nevertheless, it would be wrong to say that everything that is rule-based in composing is merely the composer's own ideas. As an indication of the opposite, it may suffice to say that in music studies, much more so than in art studies, a huge body of theory continues to be taught to all students. Rhythms, harmonic sequences, larger musical structures - all this still forms the theoretical background for what composers do, or what they try to set themselves apart from. Mandala Music attempts to capture something of this in its rules, at the cost of a certain limitation of possibilities, but with the benefit of offering an introduction to composing for amateurs.

2 DESCRIPTION OF THE GAME LEVEL AND THE COMPOSITION LEVEL IN MANDALA MUSIC

The rules of Mandala Music are as follows: Enter outlines of squares consisting of the digits 1 to 9 into the square matrix initially filled with zeros. The sequence of digits within these squares should consist of a fixed sequence of digits that is repeated several times in full. In this sequence, the same digit should never directly follow each other several times, see example in Figure 1.

The higher goal in mandala music is to try to get such squares to overlap, so that parts of the sequence of numbers in one are used simultaneously in another (interlacing), see Figure 2.

A square with an edge length of N matrix places has a length L of L=(N-1)*4 matrix places. The possible period lengths Pi of the digit sequences that can be entered in it correspond to all real divisors of L that are greater than or equal to 2 and smaller than L. For example, N=4 results in a length L=(4-1)*3=12 and therefore the periods Pi = {2, 3, 4, 6}.



0	0	0	0	0	0	0	0
0	0	5	4	5	3	2	0
0	0	6	0	0	0	1	0
0	0	2	0	0	0	2	0
0	0	1	0	0	0	6	0
0	0	2	3	5	4	5	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

FIGURE 1. FORMING A SQUARE OUTLINE OF DIGITS (CYCLE) ACCORDING TO THE RULES: STARTING FROM THE TOP LEFT AND RUNNING ANTI-CLOCKWISE, THE SAME SEQUENCE OF DIGITS IS PRODUCED TWICE IN SUCCESSION, NAMELY: 5,6,2,1,2,3,5,4. THE MATRIX SIZE OF 8X8 CHOSEN IN THE EXAMPLE SERVES AS AN EXPLANATION AND IS JUST SUFFICIENT TO PRODUCE A MUSICAL PHRASE. LARGER COMPOSITIONS REQUIRE CONSIDERABLY LARGER MATRICES

0	0	0	0	9	2	4	8
2	7	5	4	5	3	2	3
6	0	6	0	3	0	1	5
5	7	2	0	8	4	2	9
0	0	1	0	0	0	6	0
0	0	2	3	5	4	5	0
0	0	3	2	0	0	0	0
0	0	0	0	0	0	0	0

FIGURE 2. THREE FURTHER CYCLES HAVE BEEN ADDED HERE USING PARTS OF THE CYCLE FROM FIGURE 1

2.1 SYMBOLS INSTEAD OF NUMBERS AND MANDALA MODE

The numbers in the game are displayed as mirror-symmetrical symbols. This is done to improve the visual aesthetics and to make the rhythmic sequences more eye-catching. Another reason for this is to ensure that the game board can be easily viewed from any direction. And finally, it is also done because the numbers do not have to be reckoned with in the game, see Figure 3.

Finally, there is the option of mirroring the current figure twice in order to obtain something that comes a little closer to what you might imagine a mandala to look like, see Figure 4.

2.2 REWARDING INSTEAD OF PUNISHING AND FREEDOM INSTEAD OF CONTROL

The implementation of the game, see

http://www.kramann.info/25_UbiMus/02_Mandala_Music) supports the pursuit of the objectives of the game stated above, i.e. the creation of valid cycles that are as interlinked as possible, by highlighting existing cycles when you move the mouse over them. Furthermore, the music is created from the set of complete cycles. So you only get to hear something if you actually use Mandala Music to create correct cycles.

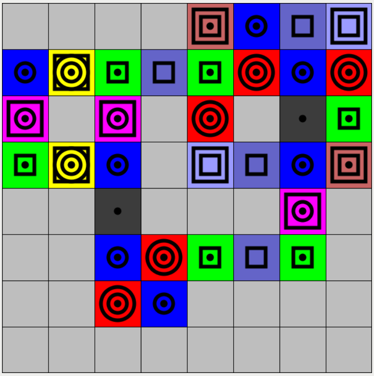


FIGURE 3. ILLUSTRATION OF THE SAME ARRANGEMENT AS IN FIGURE 2, BUT USING SYMBOLS

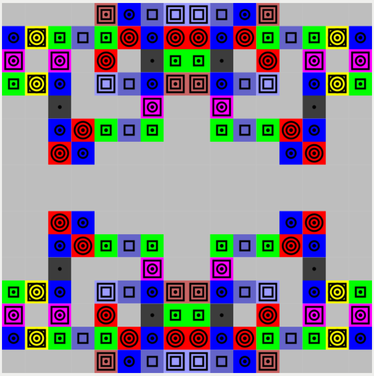


FIGURE 4. CORRESPONDS TO THE ARRANGEMENT IN FIGURE 3, BUT HERE IT IS SHOWN MIRRORED TWICE

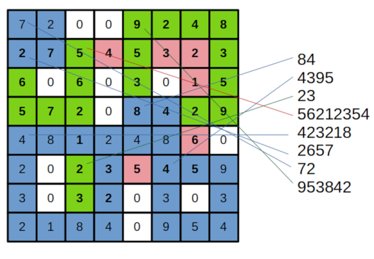
Adherence to the principle of interlacing in turn ensures that the resulting musical phrases in the voices are similar to each other. If this is accepted as a criterion for the quality of the resulting piece of music, it can be said that the resulting music is better in this sense when efforts are made to interweave. In addition, variety in the composition of the number sequences of the phrases and their length generally also ensures music that sounds versatile. However, we deliberately refrained from quantifying the named criteria and implementing them as a measure of quality for the resulting composition. This is because the user should be free to choose his own goals. The only orientation for the design of the creative process should be the aesthetic judgment of the user with regard to both the visual form in the matrix and its sonic implementation.

3 FROM MANDALA TO MUSIC

> "The Tao produced One; One produced Two; Two produced Three; Three produced All things..." (Chapter 42 from the Tao Teh King [Laozi and Legge 1891]).

The game algorithm constantly searches for all existing cycles in real time, starting from the center of the matrix and moving outwards in a clockwise spiral. Initially, exactly one period of each sequence of digits found is displayed one below the other in the list, see Figure 5.

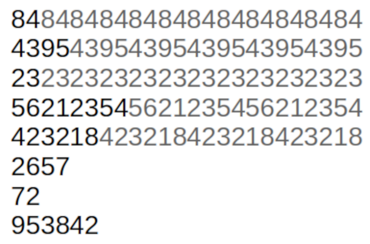
If the desired number of voices V and E is the number of entries that are to be used per voice in a time step to form a sound event, the first M=V+E-1 lines of the list are used first. The smallest common multiple K of the M lines is formed and all periods are repeated until each of the lines has the length K, see Figure 6.



7	2	0	0	9	2	4	8
2	7	5	4	5	3	2	3
6	0	6	0	3	0	1	5
5	7	2	0	8	4	2	9
4	8	1	2	4	8	6	0
2	0	2	3	5	4	5	9
3	0	3	2	0	3	0	3
2	1	8	4	0	9	5	4

84
4395
23
56212354
423218
2657
72
953842

FIGURE 5. AS FIGURE 2, BUT WITH THE ADDITION OF FOUR MORE CYCLES. NEXT TO IT: LIST OF THE NUMERICAL SEQUENCES OF THE PERIODS OF ALL EIGHT CYCLES FOUND BY THE ALGORITHM IN THE MATRIX FIGURE



848484848484848484848484
439543954395439543954395
232323232323232323232323
562123545621235456212354
423218423218423218423218
2657
72
953842

FIGURE 6. LIST FROM FIGURE 5. V=3, E=3, M=3+3-1=5 (SEE TEXT). SMALLEST COMMON MULTIPLE OF THE FIRST FIVE LINES K=24. ALL PERIODS OF THE FIRST FIVE LINES ARE REPEATED UNTIL THE LENGTH K IS REACHED

Now look at the top E numbers in each column and multiply them together. This product P, multiplied by a fixed factor F, represents the frequency frq=P*F, from which the nearest midi tone is then formed and – if within a previously defined range – played by the musical instrument assigned to this voice. Proceed in exactly the same way with the other voices, moving down one row in each case and then multiplying the next E values with each other from there. Moving one column to the right means calculating the notes to be played in the next time step. Once all K time steps have been processed, you move down just one row again and from there take M=V+E-1 rows from the list again and proceed in the same way, see Figure 7.

Because the tones for the individual voices are only ever moved down one line, the change in pitch from the tone of one voice to that of the next is only determined by the division of a small number from 1 to 9 - namely the one in the line that has been left and the multiplication of a small number from 1 to 9 - namely the one in the new line that has been added. This ensures that the frequencies of the tones are in simple numerical relationships to each other.

The harmonic and rhythmic structure of the resulting music also changes in small steps. This is ensured on the one hand by the cyclical character of the lines, but also by the fact that as the composition progresses after reaching time step K, it only moves down one line, see the musical result of the example described here: http://www.kramann.info/25_UbiMus/02_Mandala_Music/ mandalamusic.mp3 and also a second small example in Figure 8 and its corresponding sonification here: http://www.kramann.info/25_UbiMus/02_Mandala_Music/ mandalamusic2.mp3.

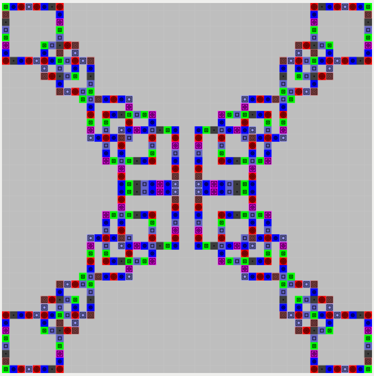


FIGURE 8. MANDALA REPRESENTATION OF A SECOND SMALL EXAMPLE

4 WORKSHOP AND INSTALLATION

Mandala music can be realized as an installation by creating as many voices (parameter V) as there are available sound sources. By assigning voices that are close to each other in terms of sequence to sound sources that are close to each other, there is a smooth transition to which voices are combined with each other when walking through the corresponding room. In the cloister of a monastery, the sound source could be positioned every ten meters to another of, for example, 12 voices, whereby the voices would then lie one behind the other in their natural sequence, i.e. 1,2,3,...12, see Figure 9.

Wiring the sound sources from an amplifier could be a problem in terms of cost and aesthetics. It would be better if the sound sources were designed as autonomous but synchronized modules. There are two options here: To use twelve laptops each with an active loudspeaker and to realize the synchronization via W-LAN, or to transfer the sequence of the cyclically repeated midi commands to microcontroller modules, which in turn control Mdi sound modules connected to active loudspeakers and to start the microcontrollers synchronously once at the beginning by means of a reset button connected to all microcontrollers in a temporary manner. The latter variant would be the most cost intensive, as less available standard components would be used, but would probably lead to the most aesthetically pleasing result.

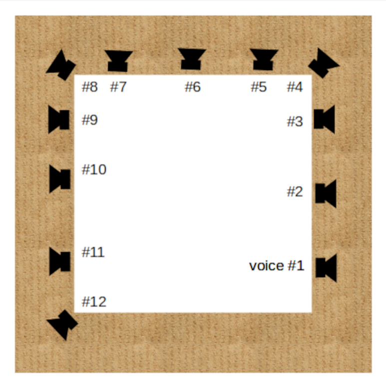


FIGURE 9. EXAMPLE OF THE ARRANGEMENT OF A SERIES OF 12 SOUND SOURCES, EACH FOR ONE VOICE OF A MANDALA MUSIC INSTALLATION IN THE CLOISTER OF A MONASTERY

The corresponding mandala matrix could take place wi-thin a workshop in which interested laypersons work together to create the mandala on a large touchscreen, ideally placed on a table around which the participants are grouped.

All the components mentioned in the description above would be provided by the author if necessary and could also be set up on site by the participants themselves.

5 IMPROVEMENT OF EVERYDAY USABILITY

It has not yet been possible to conduct an extensive user survey. However, the current experimental version of the mandala music app was made available to three people of different ages and genders (female 12, male 18, male 60) to try out, in order to obtain at least some feedback on the app's suitability and then incorporate this into the further development process.

The twelve-year-old female participant was unable or unwilling to understand the overall concept and use the app properly. Therefore, in future, I will start with young adults as the target age group.

Like the sixty-year-old, the eighteen-year-old male participant also found the game concept appealing, but complained that the music resulting from the transformation process was only updated after pressing the corresponding button. This was immediately incorporated into further development: users can now activate continuous real-time updating. This means that the acoustic equivalent of every action on the mandala is immediately perceptible.

The sixty-year-old participant could imagine spending more time with Mandala Music, the eighteen-year-old participant was unsure about this, while the twelve-year-old participant could not imagine doing so.

In contrast, the eighteen-year-old participant (see session: https://youtu.be/o44dH6uGSqY) stated that he could discern a vague connection between the current mandala structure and the accompanying music, while the other participants denied this.

The ability to switch between different views, especially between mandala and piano roll, was viewed positively by the two older test subjects. The 60-year-old test subject felt that the display on a smartphone was much too small and would have preferred a larger device as a basis. In fact, the practical implementation now uses a giant Androidbased tablet. In addition, the sound quality was improved in the practical implementation by using four-channel sound and professional physical modeling software on a laptop, Figure 10.

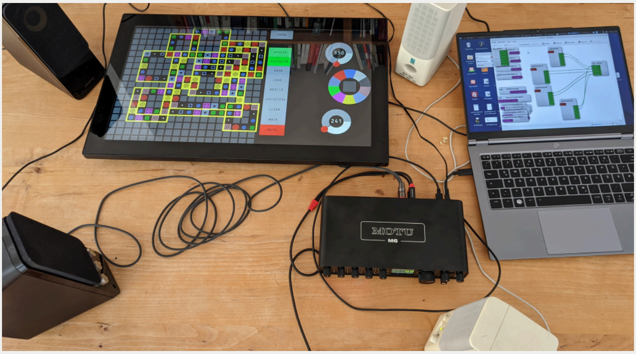


FIGURE 10. TEST SETUP WITH QUADRAPHONIC SOUND, TRANSMISSION OF MIDI DATA VIA WI-FI FROM TABLET TO LAPTOP, AND IMPROVED SOUND QUALITY THROUGH THE USE OF PHYSICAL MODELING SOFTWARE

In addition to the improved sound quality, the next version of Mandala Music will also feature impressive musical compositions created using the system, which should increase motivation to explore Mandala Music. One example is this piece for clarinet and piano entitled “Pan,” which was created using Mandala Music: http://www.kramann.info/99_Musik/alg_Pan_clarinet_piano_082025_kramann.mp3,http://www.kramann.info/99_Musik/alg_Pan_clarinet_piano_082025_kramann.pdf,https://youtu.be/ze9IG4i_Ahg,https://youtu.be/N7yhh_-zvuE, Figure 11.

6 AN ATTEMPT TO RELATE “MANDALA MUSIC” TO THE RITUAL ASSOCIATED WITH THE KALACHAKRA MANDALA

In the ritual associated with the Kalachakra mandala, monks in the tradition of tantric Buddhism painstakingly create a finely structured mandala measuring around two meters in diameter from coloured sand over a period of several days. Just one day after completion, it is destroyed again by sweeping up the sand and pouring it into a richly decorated bag. This bag is then carried in a procession to a river and the sand is poured into the water. This ritual has since been practised by Buddhist monks in many western cities, as was organized in 2016 by the art house “Sans Titre” in Potsdam, which is not far from the location of the upcoming UbiMus Symposium in Brandenburg an der Havel https://www.youtube.com/watch?v=uBZTRRK7AQI.

The meaning of the depictions in these mandalas is highly complex and can hardly be fathomed by outsiders. An essential feature, however, is a fivefold division of the mandala, consisting of the areas in the four cardinal points and the center, based on the five elements that make up the human being according to the tantric view. Formation and subsequent destruction are obviously connected with birth and death, but also with the Buddhist insight into the lack of essence behind the phenomena in which the world shows itself to us, i.e. the realization that after the removal of the structures of understanding that are valid for us, only emptiness would remain, see e.g. [Brauen 1995].

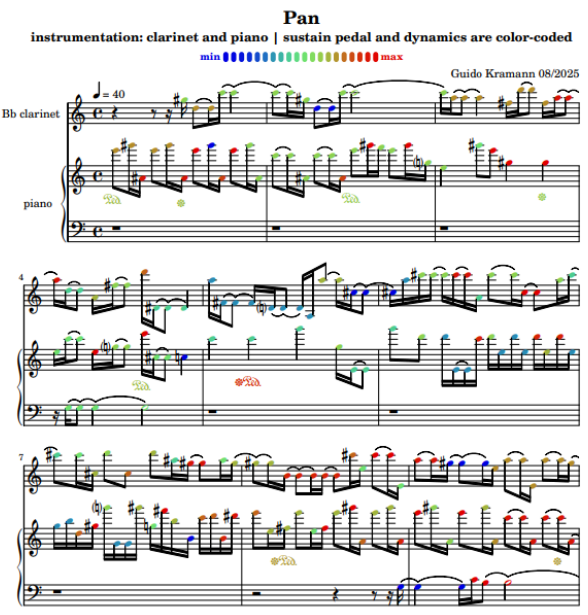


FIGURE 11. EXCERPT FROM THE SCORE OF THE COMPOSITION ”PAN” CREATED WITH MANDALA MUSIC

Of course, an artistic performance can never come close to the depth of meaning that the Buddhist mandala ritual possesses, as the former is based on a millennia-old tradition and the latter merely corresponds to the unique realization of an individual, creative idea. However, in the case of “Mandala Music as a Happening”, the reference to the external form of the traditional ritual can lead to the people involved being able to engage better and more deeply with what is happening on the spot, so that they are more involved. The parallel, that here and there a group is busy creating the mandala over a longer period of time, suggests to the visitor that it is something significant, rather than if the mandala were simply finished in front of her or him, as here the effort it takes to create it can be experienced. The parallel that all visitors are made aware that the mandala will be destroyed again at the end makes it clear that it is urgently necessary to get involved in the performance now, at this moment, because there will be no more opportunity to do so later.

7 DISCUSSION

The concept for cooperative composing presented here is characterized by a great openness to results, since although certain rules specify how tonally effective structures should be designed in the matrix, it does not specify which particular characteristics the large form of the mandala should have. It is to be expected that in the course of continued work on mandala music, users will develop a certain intuitive knowledge. Over time, you will learn that the use of the 7 tends to produce somewhat dissonant sounds. It will become clear that the combination of period durations that correspond to prime numbers in terms of length leads to a greater variety of different musical phrases and harmonies than the choice of lengths that are in simpler relationships to each other. And you will notice that the choice of the factor F, which is applied before the transformation into a miditone, decisively determines the tonal character, but also whether anything sounds at all.

Mandala Music joins the long line of contributions to Ubiquitous Music that deal with daily creative practice, called ”little c” [Keller et al. 2014]. A certain advance over previous approaches is that the approach to composing is intuitive, but offers the practical opportunity and time to make progress in composing. The simple representation of the results as a matrix of digits ensures that what has been compiled is permanently available and also allows results to be easily exchanged between users. Finally, the entire concept is so simple and completely transparent that it invites other developers to create a wide variety of implementations.

At this point it could be excusably remarked that the simplicity of the whole concept would mean that not all aspects of music creation could be incorporated into the creative framework of Mandala Music, in that although the arrangement of musical motifs with fixed pitches can be shaped within a larger rhythmic and harmonic overall structure, no other musical parameters such as dynamics, timbre, articulation or spatial sound can be influenced. However, both the simplicity of the concept and the focus on motifs, rhythm, harmony and musical form have been carefully considered. This focus was based on the conviction that it is the right one for an application intended for laypeople. In order to be able to master Mandala Music intuitively, the entire concept follows the basic idea of docking onto what is generally known and familiar. The docking to the familiar begins with the expectations that the name “Mandala Music” evokes as something that is reminiscent of coloring books for adults and that can be done on the side to relax at home or while traveling, continues with the relative similarity to sudoku and ends with the musical result that docks to the music with defined rhythms and pitches in a tempered mood that surrounds us every day. But docking onto something does not necessarily mean that it is the same as what is being docked onto. Mandala music leads the user from the familiar to something new. What is important about this sentence is: it LEADS to something new, and not: it CONFRONTS the user with something new. In this endeavour lies the deeper reason for all the conventional elements that characterize Mandala Music.

On the other hand, this transparent description of this music theory-free approach to composing may also offer others the opportunity to revisit it themselves in order to make something else out of it, which does not necessarily follow the same paradigms as the current version. Finally, it should be mentioned that in the current implementation for piano sound, the control over velocity, tempo and sustain pedal automatically adapts to the progression of the musical shape. To put it another way, in Mandala Music the unaffected musical parameters are a private matter for the user.

8 FURTHER WORK

In addition to improvements to the setup and implementation of Mandala Music in order to improve the user-friendliness of cooperative composition, it also makes sense to explore the possibility space of Mandala Music by means of search and optimization processes. Specifically, evolutionary algorithms could be written and tested that produce mandala matrices of a given size and with given properties. In this regard, there may be an opportunity to adapt the Ant Colony algorithm to Mandala Music, because as in the most prominent application of this optimization algorithm, the traveling salesman problem, paths also play an important role in Mandala Music [Dorigo and Stützle 2004]. On the one hand, this can also serve to assess what could be achieved by user groups. On the other hand, such automatically generated mandala matrices can be made available in a gallery so that users can use them as a starting point from which new mandala matrices can be created and tested by making modifications. Finally, a distributed system consisting of several software agents could be created, which try to find strategies for the cooperative creation of Mandala matrices. This can be used to draw conclusions about a meaningful design for human users working cooperatively on a mandala, but may also provide interesting results in itself, for example in a scaling that is far removed from what humans can achieve and thus generate a special virtual ecology of a creative process, see [Lazzarini et al. 2021]. And finally, one could also go in the opposite direction and instead of proceeding with the virtualization of the whole, develop a physically functioning version with a game board and game pieces, or be inspired by the ritual character of Buddhist practice when further elaborating the design of the cooperative creation of a mandala, see e.g. [Bedeaux 2024].

Just as there are amateur choirs, pottery classes, chess clubs, and the like, techniques such as those exemplified by Mandala Music could perhaps one day form the basis for something like composition clubs. As in the other examples mentioned, this would be an opportunity to immerse oneself deeply in an activity and experience a coherence between the environment and inner experience that represents a counterpoint to an increasingly fractalized world. Or, in the words of philosopher Matthew B. Crawford:

> “Skilled practices serve as an anchor to the world beyond one’s head – a point of triangulation with objects and other people who have a reality of their own” (Crawford 2015).

REFERENCES

Bedeaux, R. (2024). The tibetan sand mandala: A short history. Minneapolis Institute of Art. url: https://new.artsmia.org/stories/the-tibetan-sand-mandala-a-short-history.

Brauen, M. (1995). Das mandala - der heilige kreis im tantrischen buddhismus en: The mandala - the sacred circle in tantric buddhism. page 16. DuMont, Köln.

Crawford, M. (2015). The world beyond your head – on becoming an individual in an age of distraction. page X. Farrar, Strauss and Giroux, New York.

Cross, L. (1999). Reunion: John cage, marcel duchamp, electronic music and chess. John Cage Trust. url: https://www.johncage.org/blog/cross reunion.pdf.

Dorigo, M. and Stützle, T. (2004). Aco algorithms for the tsp. In Ant Colony Optimization, pages 67–69. The MIT Press, Cambridge.

Gottstein, B. (2024). Der klang der gegenwart (en: The sound of the present). Reclam, Regensburg.

Hamilton, R. (2014). Musical sonification of avatar physiologies, virtual flight and gesture. In Aramaki, M., Derrien, O., Kronland-Martinet, R., and Ystad, S., editors, CMMR 2013, Sound, Music, and Motion, pages 517—-532. Springer, Heidelberg.

Keller, D., Lazzarini, V., and Pimenta, M. (2014). Ubiquitous music. pages 29–30. Springer, Heidelberg.

Laozi and Legge, J. (1891). Tao teh king. Wikisource. url:
https://en.wikisource.org/wiki/T

Lazzarini, V., Keller, D., Otero, N., and Turchet, L. (2021). Ubiquitous music ecologies. pages 1–22. Routledge, London.

Merleau-Ponty, M. and Schroeder, J. (1996). Das primat der wahrnehmung. trans. jürgen schröder. 5th ed. frankfurt am main: Suhrkamp, 2016. print. trans. of le primat de la perception et ses consequences philosophiques. Éditions Verdier, Paris.

Zbyszynski, M., Donato, B. D., Visi, F., and Tanaka, A. (2021). Musical sonification of avatar physiologies, virtual flight and gesture. In Kronland-Martinet, R., Ystad, S., and Aramaki, M., editors, CMMR 2019, Perception, Representations, Image, Sound, Music, pages 600—-622. Springer, Heidelberg.

MIDI ADAPTIVE TUNING STRATEGIES BY MEANS OF AI-BASED STRUCK-STRING INTERACTION IN UBIMUS

RUI SU[1] · **JOSEPH TIMONEY**[2] · **DAMIÁN KELLER**[3]

RUI.SU.2022@MUMAIL.IE - JTIMONEY@MU.IE - DKELLER@CCRMA.STANFORD.EDU

ABSTRACT

The struck-string interaction framework provides fresh opportunities for exploration of adaptive approaches to tuning. A particularly interesting intersection of ubimus developments is the convergence of tuning with timbre-based interaction. Our analysis highlights that whereas conventional tuning techniques offer a strong theoretical and practical basis, the addition of adaptive technologies has the potential to increase flexibility. Therefore, we engage with two areas of focus within struck-string interaction: the ubimus networked-based infrastructure and the expanded notions of inharmonicity.

KEYWORDS

Struck-string interaction; tuning; adaptive techniques; inharmonicity.

1 INTRODUCTION

This paper reviews the development of tuning systems while analysing how current technology may be able to address long-standing issues of just-intonation (JI) implementations, with an emphasis on piano sounds.

---

[1] Maynooth University, Ireland.

[2] NAP, Federal University of Acre, Brazil. Ubiquitous Music Group.

We examine the state-of-the-art in piano-timbre technologies, from physical modelling to sampling, to probe the potential and constraints of fixed-pitch tuning architectures. We discuss how Just Intonation, long thought to be impracticable, may be enabled within ubiquitous music environments by fusing historical philosophy with adaptive computational techniques. After discussing recent developments in AI-driven adaptive tuning, we speculate on a web-based JI plugin design that uses chord analysis and real-time pitch recognition.

The widespread usage of 12-tone Equal Temperament (12-TET) in the late Renaissance and Baroque was a result of centuries of tuning system innovations and compromises. In contrast to previous systems and fostering the adoption of tonality, 12-TET ensured uniform intervals across all keys by dividing the octave into twelve equal logarithmic steps. This homogeneity impacted pitch-based compositional practices enabling easier key modulations when compared to Pythagorean and Meantone tuning systems (Stange et al., 2018). 12-TET deviates slightly from natural frequency ratios. While guaranteeing consistent modulations, the 12-TET approach features slight dissonances dispersed across the scale, sacrificing the clean-sounding, beat-free intervals of Just Intonation. In contrast, JI produces harmonic series-aligned intervals that have a warm, resonant quality but it rapidly becomes unwieldy in multi-key settings. This diversity of approaches to tuning suggests a potential conflict between practicality and purity, exemplified by the adoption of equal temperament as an alternative to just intonation, that has not been resolved[3].

Many current musical practices use tunings based on natural intervals, such as the systems based on Chinese pentatonic scales which prioritise simplicity and resonance, and the Indian Shrutis, which divide the octave into 22 microtonal steps. Adaptive tuning algorithms may foster new paths for developing legacy tuning systems that so far have been constrained by fixed-pitch instruments.

The remaining sections examine how tuning is being reconsidered through the application of artificial-intelligence and digital signal processing tools, in an effort to find adaptive solutions. The first section explains prior work on tuning systems and discusses their advantages and limitations, including alternative tunings, dynamic adaptive intonation, and just intonation algorithms. The second section explores the technical requirements for the development of struck-string interaction prototypes, highlighting network-based infrastructure and expanded inharmonicity as areas of focus. We then present musical examples and provide pointers to future design efforts in struck-string interaction.

JUST INTONATION

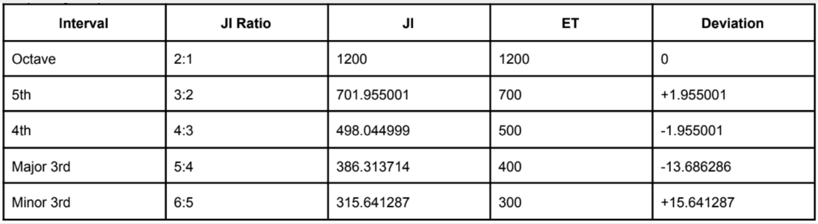
Just Intonation is a tuning method based on straightforward rational frequency ratios (see also Hermode Tuning below). This approach may be summarised as

---

3. Our approach to tackling multiple tuning systems within a consistent framework stems from a ubimus-design perspective of expanding the aesthetic pliability of music-making. Similarly to meter-based temporalities, current tuning techniques restrict the range of sonic organisations to a subset of preset options.

continually solving a system of linear equations, as opposed to making a series of if-then decisions. The set of equations can be thought of as a mechanical network of springs or resistors representing the interval sizes. Each spring will, whenever feasible, tend toward a state where its length matches the size of a pure interval. If the spring network is very intricate and it is not feasible for every spring to be in its ground state, the system may approach a non-trivial state under tension (e.g., a tempered harmonic compromise). This may be compared to how musicians discover in-tune intervals. A benefit of this approach is that it finds a balanced set of all the targeted intervals, not just nearby tones (Stange et al., 2018).

TABLE 1. COMPARATIVE ANALYSIS OF JUST INTONATION (JI) AND EQUAL TEMPERAMENT (ET) INTERVAL VALUES. THIS TABLE DISPLAYS VALUES FOR PREVALENT MUSICAL INTERVALS ACCORDING TO JUST INTONATION RATIOS AND EQUAL TEMPERAMENT, INCLUDING THE VARIANCES IN CENTS. IT ILLUSTRATES THE DISTINCTIONS BETWEEN THE TWO TUNING SYSTEMS, EMPHASISING THE SMALL FREQUENCY DISPARITIES INHERENT IN EACH METHOD



Interval	JI Ratio	JI	ET	Deviation
Octave	2:1	1200	1200	0
5th	3:2	701.955001	700	+1.955001
4th	4:3	498.044999	500	-1.955001
Major 3rd	5:4	386.313714	400	-13.686286
Minor 3rd	6:5	315.641287	300	+15.641287

In contrast to other tuning techniques, just intonation uses vertical and horizontal harmony to tune the pitch content. Adaptive tuning methods face two significant tuning aspects, as pointed out by Stange et al. (2018). On the one hand, every new chord must be intoned ‘vertically,’ meaning the relative pitches are performed simultaneously. However, the harmonic progression requires the following chords to be adjusted in relation to one another according to ‘horizontal’ (or sequential) criteria[4]. In just intonation, the twelve pitch classes are adjusted so that, with respect to a certain reference frequency f∗, all frequency ratios can be expressed as simple rational values. The steady migration of the overall pitch is one of the main drawbacks of dynamic tuning methods that rely on sequential methods.

Stange, Wick, and Hinrichsen (2018) describe an implementation of JI that features built-in microtonal sound generators and real-time MIDI instrument modifications. Its support of dynamic polyphony and pitch-bend commands enables experiments with adaptive tuning and comparisons of the outcomes of JI with ET and other static tunings.

Another method, based on the psychoacoustic concept of dissonance, was proposed by Sethares (2002). In order to establish a cost function for a gradient-based optimisation process, the system continuously responds to the changes in JI ratios. By lowering the expense parameter, the pitches are automatically adjusted to reduce dissonance. An example of how a dynamic tuning model can lessen the dissonance was presented in Kirsch and coauthors’ work (Kirsch, 2021). Their algorithm minimises estimated roughness between simultaneous pitches or chords by optimising frequencies. The algorithm was then adjusted to optimise the dissonance reduction after the results were compared to conventional tuning techniques. The overall

---

[4] Combinations of tones are typically interpreted as either in tune or out of tune when performed simultaneously. But their intonation is also relevant consecutively. Smooth harmonic transitions tend to focus on the pitches of succeeding chords to account for intonational memory. Consequently, an intonational discontinuity between chords may result from an abrupt change in the chordal root if the chords are unrelated.

computed dissonance reduction averaged up to 12%.

ALTERNATIVE TUNINGS

Kite Giedraitis' Alt-tuner was created as a plug-in for digital audio workstations (DAWs) to offer a versatile platform and dynamic pitch adjustment for recordings and live performances based on the MIDI protocol (Giedraitis et al., 2019). Particularly in JI, Alt-tuner's ability to adjust tunings on the fly provides support for group performances in multi-keyboard setups. Alt-tuner offers meantone, various equal temperaments, and microtonal scale tunings in addition to just intonation tuning. Additionally, it features a user-defined tuning system that lets users create their own scales.

Nevertheless, Alt-tuner has some drawbacks due to a complicated procedure that necessitates knowledge of tuning systems, microtonal ideas, and just intonation ratios. Thus, its custom settings can be difficult for novices. Given that not all DAWs natively support microtonal tuning, its reliance on a DAW is also a caveat.

HERMODE TUNING

Hermode Tuning (HMT) is a method for dynamically tuning virtual instruments (Mohrlok, 2003). The implementation is predicated on the notion of examining harmonic structures and adjusting pitches to achieve a consistent intonation. To maintain compatibility with well-known tools, HMT also supports re-tuning based on equal temperament.

Over time, a number of tuning techniques have been developed; Hermode Tuning is among the well-known adaptive tuning systems, with applications ranging from plug-ins for software programs like Cubase to implementations in church organs. Rather than figuring out the chordal root, the algorithm adjusts intervals between the vertically adjacent tones of a particular chord to fit specific ratios. To minimise deviations from the standard equal-temperament, the pitch of the fundamental is simultaneously modified. As a result, this reduces the frequency shifts between subsequent chords (Stange et al., 2018). Figure 1 shows how HMT tunes A major chords and their resulting differential tones.

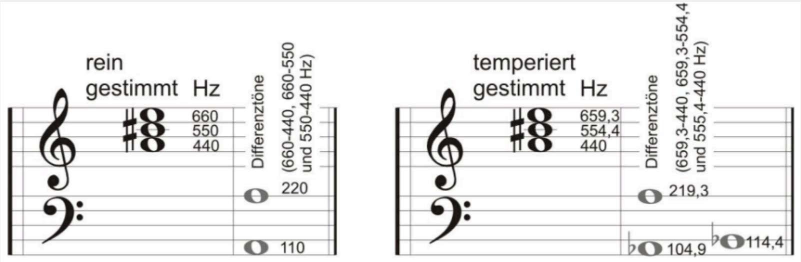


FIGURE 1. 12-TET A MAJOR CHORD VS. HMT-TUNED A MAJOR CHORD (MOHRLOK, 2003)

Hermode tuning allows selecting a preferred tuning style. Distinct tuning modes prioritise just intonation. All HMT modes, for instance, restrict the effective line of equal temperament's variation from a baseline of +/-20 cents. This indicates that no frequency can vary from the level of equal temperament by more than +/-30 cents. The natural seventh in the HMT mode Jazz/Pop is the only exception, with a largest deviation set to -47 cents.

In contrast to previous adaptive tuning techniques, the concept of harmonic centre guarantees that tuning stays constant across key changes, thus avoiding sudden shifts.

The current harmonic centre is determined by analysing the last 10 chords and saving the computed tuning values in memory. Only the tuning values of the final ten chord structures remain, substituting the first, second, etc. from the ongoing analysis as a new harmonic centre emerges.

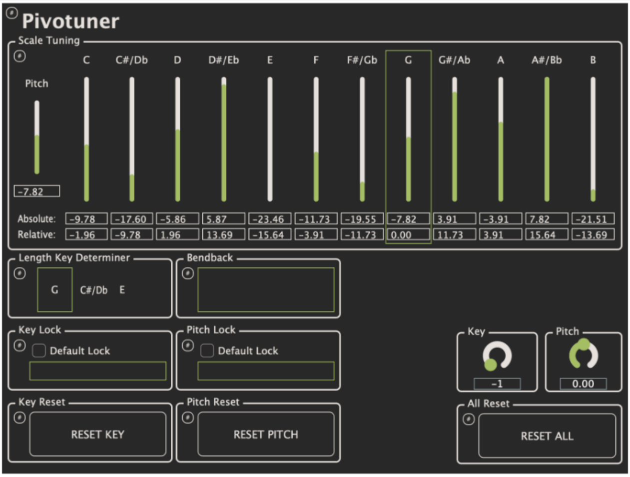
HMT applies just intonation modifications through smooth harmonic transitions but it may present certain drawbacks: due to its heavy reliance on equal temperament and just intonation, HMT might not be able to handle alternative scales; the majority of DAWs do not support it; furthermore, in highly chromatic music where harmonies change quickly, it may not have enough information to establish a definite harmonic centre.

PIVOTUNER

The VST3/AU MIDI plugin Pivotuner automatically adjusts pitch data to provide support for adaptive pure intonation (Volkov, 2023). Finding a reference key is a first step: Pivotuner employs a Key Determiner method to identify a reference note to adjust the pitches of a MIDI data stream. After identifying a key, Pivotuner determines how to tune the subsequent notes. Provided that non-equal temperament intervals are specified, Anytime the Key Determiner reports a new key, a microtonal modulation takes place, (Volkov, 2023). The Pivotuner plugin interface is featured in figure 2.

Then, using an adjustable-tuning algorithm, Pivotuner adjusts the active pitches in relation to the given key. In addition to choosing the algorithm for key and tuning selection. It also provides ways to regulate the settings when a new key is selected, to control microtonal modulation, and to tune each interval and chord individually.

Pivotuner does have some limitations, though. For intervals from minor 2nd to major 7th, Pivotuner allows the use of any tuning; however, up to this point, only ratio-based pure tunings have been employed. Additionally, Pivotuner has not been utilised in live settings with voice or acoustic instruments. Although performing microtonal music with voice or on acoustic instruments is challenging, the proposed technique might be able to supply microtonal reference pitches to make this approach deployable in performances (Volkov, 2023).



Pitch: -7.82

	C	C#/Db	D	D#/Eb	E	F	F#/Gb	G	G#/Ab	A	A#/Bb	B
Absolute:	-9.78	-17.60	-5.86	5.87	-23.46	-11.73	-19.55	-7.82	3.91	-3.91	7.82	-21.51
Relative:	-1.96	-9.78	1.96	13.69	-15.64	-3.91	-11.73	0.00	11.73	3.91	15.64	-13.69

Key: -1

Pitch: 0.00

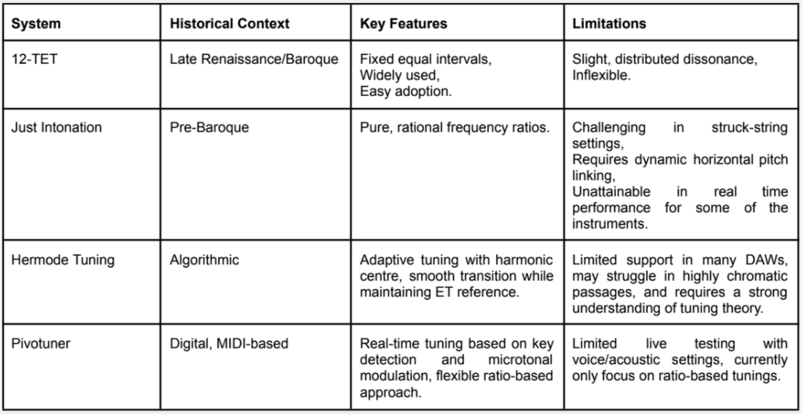
FIGURE 2. PIVOTUNER USER INTERFACE (VOLKOV, 2023)

MULTI-PITCH ESTIMATION

The research team from Deep Autotuner created a machine learning algorithm that uses instrumental accompaniment recordings as references to predict pitch correction for a monophonic vocal track (Wager et al., 2020). Additionally, their findings on a CNN with a GRU layer show that the amount of pitch correction may be determined with the help of spectral information extracted from the vocal and accompaniment tracks.

Deep learning techniques based on convolutional architectures and HCQT input constitute an efficient and reliable approach to general-purpose MPE (Multi-Pitch Estimation). According to Weis and Peeters's (2022) comparison of various architectures based on convolutional neural networks, the U-net structure, and self-attention components. This approach to deep learning might handle pitch information extraction which in turn may support Just Intonation tuning implementations.

TABLE 2. COMPARATIVE ANALYSIS OF TUNING SYSTEMS. THIS TABLE COMPARES HISTORICAL CONTEXT, PRIMARY FEATURES, AND CONSTRAINTS OF FOUR TUNING SYSTEMS: 12-TONE EQUAL TEMPERAMENT (12-TET), JUST INTONATION (JI), HERMODE TUNING (HMT), AND PIVOTUNER. IT OFFERS A SUCCINCT COMPARISON TO FACILITATE COMPREHENSION OF THEIR HISTORICAL DEVELOPMENT AND TECHNICAL INTRICACIES



System	Historical Context	Key Features	Limitations
12-TET	Late Renaissance/Baroque	Fixed equal intervals, Widely used, Easy adoption.	Slight, distributed dissonance, Inflexible.
Just Intonation	Pre-Baroque	Pure, rational frequency ratios.	Challenging in struck-string settings, Requires dynamic horizontal pitch linking, Unattainable in real time performance for some of the instruments.
Hermode Tuning	Algorithmic	Adaptive tuning with harmonic centre, smooth transition while maintaining ET reference.	Limited support in many DAWs, may struggle in highly chromatic passages, and requires a strong understanding of tuning theory.
Pivotuner	Digital, MIDI-based	Real-time tuning based on key detection and microtonal modulation, flexible ratio-based approach.	Limited live testing with voice/acoustic settings, currently only focus on ratio-based tunings.

This section reviewed approaches to dynamic tuning, highlighting caveats and opportunities for deployment in the context of struck-string interaction. Tuning approaches have traditionally focused on the organisation of scales and pitch-oriented strategies strongly based on tonality. More flexible techniques may be fostered through the application of adaptive, context-sensitive techniques and a focus on timbre-based interaction. We will develop these ideas in the next sections. Having covered the recent work on tuning systems, we now turn to struck-string interaction developments. Our analysis of requirements highlights two areas of focus: network-based requirements and applications of inharmonicity.

APPLICATIONS OF ADAPTIVE TUNING STRATEGIES IN STRUCK-STRING INTERACTION: EXPANDED PIANO INHARMONICITY

When an object is struck or plucked, its fundamental frequency is the natural frequency at which it vibrates, creating the lowest and most prominent pitch. The harmonic series is made up of integer multiples of a vibrating object's lowest frequency. However, piano strings feature in-harmonic partials—resonant frequencies that differ from these harmonics. This discrepancy arises because ideal strings devoid of stiffness are assumed by the idealised harmonic series, which is never the case in practice. Inharmonicity is the term used to describe the extent of this deviation caused by the stiffness of strings. Tuning relies heavily on inharmonicity (Roy, 2024).

Fletcher arrived at the equation of motion of a string fixed at either end, taking into account the tension and elastic stiffness that produce a restoring force, as well as the energy conservation principles in the bending and stretching of the string (Rasch & Heetvelt, 1985; Roy, 2024). In equation 1, the frequency, fn, represents the nth mode of a string, where the frequencies of any two modes may be precisely measured to yield two constants, *F* and *B*. *B*, which has units of $m^{-2}s^{2}$, is the inharmonicity coefficient. The structural characteristics of piano strings, such as material stiffness and diameter-to-length ratios, cause deviations that are measured by the inharmonicity coefficient *B*.

$$f_n = nF\sqrt{(1 + Bn^2)}, \qquad \textbf{(EQUATION 1)}$$

Commonly seen in lower-register piano strings, the magnitude of B increases with shorter lengths and thicker strings (Dalmont, 2021). This effect is particularly important for lower bass strings, especially for smaller pianos where normally they would feature very thick strings relative to their length. Generally speaking, larger pianos tend to have longer strings so generally exhibit lower inharmonicity, particularly in the bass. Typically, a piano's lower tones have "wound" strings, the middle and upper tones "plain" strings. On the average, the inharmonicity of the wound strings is smaller than that of the plain strings – inharmonicity is not consistent over the keyboard; in the wound portion, it generally rises as one descends the keyboard , while in the plain wire section, inharmonicity tends to be low around the middle registers (like the octave below C4) and increases towards the higher notes (Rasch & Heetvelt, 1985). The

Railsback curve, which shows variations from equal temperament tuning across the keyboard, has been widely modelled and connected with stretched tuning, a typical aspect of piano tuning (Jaatinen and Pätynen, 2022; Rasch and Heetvelt, 1985).

Piano tones' inharmonicity results in a tuning curve quite different from its idealised harmonic counterpart and resembles the Railsback curve. For aural piano-tuning, beats are essential. In other words, the inharmonicity of the strings has a major impact on the tuning stretch. According to Jaatinen and Pätynen (2022), most of the inharmonics are heard as higher in the low and middle registers (up to C#7). Uncertainty tones and no discernible patterns are featured in the highest register (above C#7). Subjective octave experiment's findings are in line with those of professional musicians, despite the fact that some of the tuners have no prior experience performing music. Inharmonicity may no longer play a significant role in correcting pitch recognition because the majority of harmonics in the highest register reach the human pitch detection limits. Thus, the subjective octave-based approach replaces the beat-based approach at the highest register (Jaatinen and Pätynen, 2022).

Inharmonicity has been the subject of several investigations involving materials. Optimising string winding and shape to balance stiffness and harmonicity is one example of a construction method. Recent developments concentrate on analytical and experimental material techniques, like adding mass close to the string ends or using string densities that vary sinusoidally (Dalmont, 2021). Acoustic pianos could be expanded through experimental research on temperature-dependent self-tuning systems, aiming to avoid manual tuning (Roy, 2024).

Having discussed inharmonicity as a target application in struck-string interaction, we now turn to methods to enable flexible tuning techniques. We build on a ubimus tradition of recycling technological resources for innovation. Thus reducing the ecological footprint while expanding the integration of ubimus infrastructure.

UBIMUS NETWORKED-BASED SSI DEPLOYMENTS: WEB MIDI AND MIDI 2.0

MIDI (Musical Instrument Digital Interface), a well-known technical standard, is still widely used in the fields of electronic music creation and musical information retrieval even after four decades of use. A recent development is MIDI 2.0, made available in the early 2020s. MIDI 2.0 is an update to the MIDI specification that aims to address part of the shortcomings of the original MIDI 1.0 protocol. In addition to providing space in the specification for upcoming advances, MIDI 2.0 seeks to increase the range of tools while maintaining backward compatibility with the extant MIDI 1.0 infrastructure.

MIDI 2.0 has features relevant to adaptive tuning and struck-string interaction. By including techniques for two-way communication between MIDI devices using MIDI Capability Inquiry (MIDI-CI) messages to send and receive device information, it overcomes some of the shortcomings of the original MIDI standard. New features include an enlarged data format for higher resolution with extensibility targeting future definitions of messages, as well as auto-configuration by bidirectional connections that allows devices to discover details about other connected devices.

Importantly, it standardizes per-note controllers, letting parameters like pitch bend, expression, or articulation data affect individual events, instead of the whole MIDI channel. This refinement presents interesting possibilities for complex adaptive tuning techniques, including dynamic Just Intonation. Figure 4 shows a comparison between channel-based control of MIDI 1.0 and per-note capability of MIDI 2.0. Although hardware and software adoption is still in progress, some devices have already implemented MIDI 2.0.

Developed by W3C Audio Working Group, the Web MIDI application programming interface (API) aims to bring MIDI to the Web by enabling communication and interaction between a browser application and MIDI devices. One of the goals is to make MIDI devices support a standard feature of OS systems and web browsers on various hardware platforms. The aim is to allow web applications to communicate directly with MIDI-enabled devices, increasing accessibility for online music production and live performances (Baratè & Ludovico, 2022). Web applications may list, choose, and communicate with MIDI input/output devices thanks to the JavaScript interfaces that the Web MIDI API exposes. Because of this integration, software installations are no longer necessary, fostering the expansion of web-based DAWs, MIDI controllers, and interactive music applications. Web MIDI API has already achieved some level of low latency and real-time communication. However, it also faces difficulties, even though popular browsers like Google Chrome and Edge support it, platform popularity is still developing. Widespread adoption has been hampered by security concerns, especially those related to device access and potentially harmful interactions (Baratè &

Ludovico, 2022). Furthermore, support for accurate timing is limited (Keller and Lazzarini, 2024).

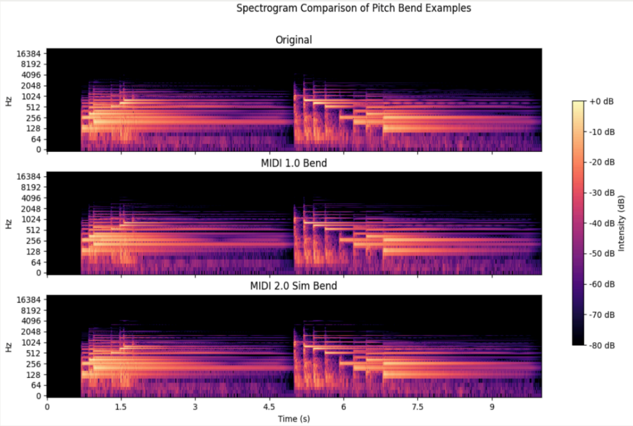


FIGURE 3. SPECTROGRAM COMPARISON ILLUSTRATING MIDI 1.0 VS. SIMULATED MIDI 2.0 PITCH CONTROL ON A C MAJOR CHORD. THE SPECTROGRAM SHOWS THE TIME VS. FREQUENCY. WITH THE TOP SCALE WE CAN SEE IT HAS THE MOST STABLE FREQUENCIES. THE MIDDLE SCALE REPRESENTS MIDI 1.0 PITCH BEND, WITH HARMONICS OF ALL NOTES CURVE SIMULTANEOUSLY (ESPECIALLY FROM 1.5S - 3S). AT THE BOTTOM SCALE, MIDI 2.0 PER NOTE SIMULATION, HARMONICS OF STABLE NOTES (C4, G4, C5) REMAIN STRAIGHT, WHILE ONLY THE HARMONICS OF THE TARGET (E4 ~ 392 HZ) CURVE, DEMONSTRATING INDEPENDENT PITCH CONTROL CRUCIAL FOR PRECISE ADAPTIVE TUNING. SAMPLE AUDIOS IN APPENDIX.

Tools like Midiano[5] feature the possibility of running MIDI apps straight in the browser using the Web MIDI API. Users of this web software can load standard MIDI files, see the events (usually in a 'piano roll' or falling-note style), link a physical MIDI unit for interactive playback and feedback, and set the playback speed.

OPEN ISSUES AND FUTURE DEVELOPMENTS

Our study of the state-of-the-art in adaptive tuning shows

---

5. https://midiano.com/

that some progress has been made in the last decades but there are still problems with accessibility, polyphonic control, and aesthetic pliability. This suggests a direction for future research: developing a comprehensive, browser-based framework that combines the emerging techniques while targeting an aesthetically flexible design. Future struck-string interaction systems should not only address long-standing issues related to Just Intonation, they may also foster new possibilities in timbre manipulation, for instance, by featuring dynamic control of piano string inharmonicity. Therefore, two areas of focus for the development of struck-string interaction are proposed: integrated development of ubimus networked-based infrastructure and the expanded notions of inharmonicity enabled by adaptive tuning.

Inharmonicity is inherent to piano strings yielding acoustic features that distinguish each instrument. Theoretical knowledge, practical engineering, and psychoacoustic considerations are necessary to balance its effects. Our understanding and use of this phenomenon will grow as research progresses thanks to a stronger connection between the physics of sound and the craft of music.

Machine-learning pitch detection and adaptive tunings may offer the subtle nuance required to replicate the retuning practices done by musicians and musicologists in their daily studies. By analysing live musical contexts and implementing per-event control capabilities based on the latest updated protocol MIDI 2.0, ubimus systems may be capable of aligning the current tuning with rapid changes of pitch-based systems. Taken together, the approaches discussed in this paper show how tuning systems and cutting-edge technologies open up new possibilities and increase the viability of using Just Intonation in both legacy and prospective ubimus frameworks.

Nevertheless, special attention needs to be given to the emergent caveats of the adoption of AI-based tools. Despite the fact that ethical and political concerns surrounding AI tools have gained more attention in recent decades, there has been a dearth of research on the effects of AI techniques on music communities. The OECD AI standards and nine well-known ethical statements were examined by Oğul (2024), pointing to a tentative convergence towards transparency, human-centered values, fairness, and privacy standards. This initiative may offer a useful framework for the development and application of responsible AI in emergent, community-based music practices.

Stricter ethical standards should be enforced. For instance, some generative systems adopted by the music industry do not take into account the environmental impact of their use, from energy consumption to hardware lifecycle. We hope that the continuous efforts of the ubimus community may eventually open the door to AI applications that work in concert with humans to create dynamic and morally sound musical practices.

Our ongoing work targeting dynamic tuning mechanisms, featuring browser-based toolkits, may tackle inharmonicity issues in piano sounds by applying predefined tuning ratios (Su et al., 2024). This approach bridges theoretical tuning models with web deployment. Building on that foundation, we are currently updating the code base to support MIDI 2.0 and AI-based adaptive techniques. Challenges such as latency, computational load, and ensuring the accuracy of pitch detection remain active areas for research. Our next steps focus on refining these issues,incorporating feedback from performance trials, and exploring the integration of emerging technologies like edge computing and cloud services.

The next important step is to validate such a system thoroughly. Testing its strengths and weaknesses involves using musical excerpts for comparative analysis with the latest state-of-the-art systems. One way to check for vertical tuning purity is to process MIDI recordings of harmonically stable pieces, like a J.S. Bach chorale. On the other hand, pieces from highly chromatic late-Romantic works or jazz excerpts that feature harmonic complexity would test the reactive algorithm's preciseness and malleability during key changes. Then, we may compare the performance of the system to tools like Pivotuner and Hermode Tuning to achieve a quantitative and qualitative benchmark.

Finally, another way to progress would be to go beyond pitch sequences and look into how adaptive tuning can be used on timbre itself. The suggested framework could be used in additive synthesis as well. Instead of just changing the fundamental pitch, the system could be designed in a way to change the inharmonicity of the partials, and to change the synthesised piano tones to match the detected harmonic context. This combination of adaptive tuning and real-time timbral synthesis is an area for artists and engineers to explore in future struck-string interaction initiatives. In conclusion, grounded on solid musical tuning theories and cutting-edge adaptive techniques, struck-string interaction tuning may open new avenues for artistic exploration.

REFERENCES

Baratè, A., & Ludovico, L. A. (2022). Web MIDI API: State of the art and future perspectives. Journal of the Audio Engineering Society, 70(11), 918–925. https://doi.org/10.17743/jaes.2022.0028

Dalmont, J. (2021). Piano bass strings with reduced inharmonicity: theory and experiments. Acta Acustica, 5, 9. https://doi.org/10.1051/aacus/2021002

Giedraitis, K., Tall Kite software, Paul Erlich, Maria Olaya, Dave Ryan, Margo Schulter, Cam Taylor, Mike Battaglia, Abhi Timbadia, Praveen Venkataramana, Aaron Wolf, & Jason Yerger. (2019). Alternative Tunings: theory, notation and practice.

Jaatinen, J., & Pätynen, J. (2022). Effect of inharmonicity on pitch perception and subjective tuning of piano tones. The Journal of the Acoustical Society of America, 152(2), 1146–1157. https://doi.org/10.1121/10.0013572

Keller, D., & Lazzarini, V. (2024). Building blocks for lite coding: Temporalities in LitePlay.js. Proceedings of the Ubiquitous Music Symposium 2024 (UbiMus2024), 32. https://www.researchgate.net/publication/389055288_Proceedings_of_the_Ubiquitous_Music_Symposium_2024_UbiMus2024

Kirsch, K. (2021). Reducing Dissonance with Dynamic Tuning Algorithms for MIDI Synthesis (By D. Gregg, University of Dublin, & School of Computer Science and Statistics, O'Reilly Institute, Trinity College) [Thesis]. https://publications.scss.tcd.ie/theses/diss/2021/TCD-SCSS-DISSERTATION-2021-065.pdf

Mohrlok, W. (2003). The Hermode Tuning system (pp. 1–3). https://sethares.engr.wisc.edu/paperspdf/hermode.pdf

Oğul, S. (2024). In Tune with Ethics: Responsible Artificial Intelligence and Music Industry. Istanbul Bilgi University. https://doi.org/10.47613/reflektif.2024.150

Rasch, R. A., & Heetvelt, V. (1985). String inharmonicity and piano tuning. Music Perception: an Interdisciplinary Journal, 3(2), 171–189. https://doi.org/10.2307/40285331

Roy, E. (2024). Investigating the inharmonicity of piano strings. Edinburgh Student Journal of Science, 1(2), 1–4. https://doi.org/10.2218/esjs.9815

Sethares, W. A. (2002). Real-Time adaptive tunings using Max. Journal of New Music Research, 31(4), 347–355. https://doi.org/10.1076/jnmr.31.4.347.14163

Stange, K., Wick, C., & Hinrichsen, H. (2018). Playing Music in Just Intonation: a dynamically adaptive tuning scheme. Computer Music Journal, 42(3), 47–62. https://doi.org/10.1162/comj_a_00478

Su, R., Timoney, J., & Keller, D. (2024). Just intonation and inharmonicity in struck-string interaction. Proceedings of the Ubiquitous Music Symposium 2024 (UbiMus2024), 44. https://www.researchgate.net/publication/389055288_Proceedings_of_the_Ubiquitous_Music_Symposium_2024_UbiMus2024

Volkov, D. (2023, June 6). Pivotuner: automatic real-time pure intonation and microtonal modulation. arXiv.org. https://arxiv.org/abs/2306.03873

Wager, S., Tzanetakis, G., Wang, C.-i., & Kim, M. (2020). Deep Autotuner: A pitch correcting network for singing performances. ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 246-250). IEEE. https://doi.org/10.1109/ICASSP40776.2020.9054308

Weis, C., & Peeters, G. (2022). Comparing deep models and evaluation strategies for Multi-Pitch estimation in music recordings. IEEE/ACM Transactions on Audio Speech and Language Processing, 30, 2814–2827. https://doi.org/10.1109/taslp.2022.3200547

APPENDIX

Audio Example 1: Reference C Major chord, no pitch bend. https://drive.google.com/file/d/1wOnWHfat7QgVhZVBnDKxCzcV2vOtHnyX/view?usp=sharing

Audio Example 2: C Major with MIDI 1.0 channel pitch bend (all notes bend simultaneously). https://drive.google.com/file/d/18Po_wOJ-HIOmStQNeYpy6bY6FcXch9qZ/view?usp=sharing

Audio Example 3: C Major simulating MIDI 2.0 per-note pitch bend (E4 bends independently). https://drive.google.com/file/d/1Rc3vF_pA9pC6BNbm3yT5F6XHNZ-ahRi4/view?usp=sharing

SECTION II - DEALING WITH ACOUSTIC-INSTRUMENTAL RESOURCES

Exploring Immersive Impact through Vertical and Horizontal Sound Design in Hybrid Orchestral Composition (Peters; Koszolko; Scott).

EXPLORING IMMERSIVE IMPACT THROUGH VERTICAL AND HORIZONTAL SOUND DESIGN IN HYBRID ORCHESTRAL COMPOSITION

RYAN PETERS, MARTIN K. KOSZOLKO, NATHAN SCOTT

RYAN.PETERS@UON.EDU.AU

ABSTRACT

Immersive audio is an aspect of music production enabling sound to be positioned around a person, creating a more engaging listener experience. The increased accessibility to immersive audio technology has created a new paradigm for composers, delivering a new tool set of creative possibilities.

Creating immersive audio content requires a broad skillset encompassing technical sound production skills, an understanding of the implications and affordances on the compositional process, and the impact of immersive sound on the listener. Without an understanding of these areas, composers are not empowered to effectively engage with immersive audio early in the compositional process, rather placing an increased reliance on later production processes to position musical material in the audio environment.

This paper explores how knowledge of immersive audio can inform and influence music composition. It reports on a new creative work Immersio Overture, which applies immersive techniques in the context of hybrid orchestral music - combining traditional and electronic instruments such as synthesisers, guitars, and

percussion ensembles.

Through a practice-based approach, the study addresses the question: How effectively can selected immersive audio techniques be applied to hybrid orchestral music through composition?

This project investigates the positioning of musical instruments on the vertical and horizontal axis, the use of room sounds and spatial placement, and the roles of additional supporting elements, in order to examine how these factors create sonic experiences distinct from stereo. It also reflects on the emotional impact that such techniques may invoke in listeners. Immersio Overture demonstrates, through the panning of instruments such as drums, how perspectives can be shifted, locations re-imagined, and ambiences or reverbs juxtaposed to shape perceptions of size, space, and ensemble configuration.

The work also considers the technical and creative ramifications of distributing layers along the height axis,along the height axis, thereby providing greater clarity and divergence between instruments without creating an audible distraction.

The findings highlight the efficacy of several immersive audio characteristics that can enhance musical composition and provide additional expressive impact.

Link to creative work, Immersio Overture (2:22): https://www.dropbox.com/scl/fo/33ba8ay1hgcnsyh6w8pog/AB8AGAHZn1grSfbVTgFz 2I?rlkey=dnrofrpu3rktu2233n31vetk8&st=yn9tptoa&dl=0

1 INTRODUCTION

Music that uses immersive audio is produced using audio systems in which speakers are positioned around the listener in multiple layers. As these technologies become increasingly accessible, the field is attracting scholarly and industry attention, enabling innovative production techniques that can create unique experiences specific to the format.

This project explores immersive impact through vertical and horizontal sound design in hybrid orchestral composition. Hybrid orchestral music uses traditional orchestral instruments combined with electronic instruments such as synthesisers, guitars, non-rhythmic or non-tonal instruments and a variety of percussion ensembles. We explore production and composition approaches that are particularly relevant to immersive audio through a composition titled Immersio Overture. This research aims to address the question: How effectively can selected immersive impact techniques be applied to hybrid orchestral music through composition? The discussed music composition is available in multiple formats, including stereo, binaural and surround sound, plus immersive sound formats Dolby Atmos ADM and ambiX. It is recommended that a 7.1.4 listening environment is used to enable the full experience of the discussed techniques and concepts.

2 COMPOSITIONAL TECHNIQUES

This study focuses on four areas as applicable to the creative work: immersive impact, vertical axis separation, non-distractive height energy and microphone distance panning.

This approach facilitates an understanding of how these techniques are creatively applied and considers the roles the techniques play in creating immersive music from its foundation.

2.1 IMMERSIVE IMPACT

The immersive impact (the sudden shift of instruments positioning, as well as their volume, role or timbre to create the impression of sounds suddenly appearing around the listener) can be achieved with both instrumental and non-instrumental sources. In titled Immersio Overture, at 00:48, footsteps are automated across channels, moving around the listener to create a vivid sense of space. These footsteps act as a transitional cue, signalling a forthcoming change and the introduction of new musical elements and locations. The height speakers located above the listener, deliver the sound of rainfall, creating a sonic presence in the elevated space. To reinforce this spatial impression, a room reverb was added, giving the listener a perspective of size, location and room materials. At 00:57, a full drum circle commences, placing the listener at the centre of the performers. Each of the ear-level speakers features a direct percussion instrument across the horizontal plane, while the height speakers produce room and reverb sounds.

This immersive impact may be understood as an extension of the "drop" – an effect in electronic music whereby "following a build-up and often a break, there is a sudden rhythmic, timbral, and/or harmonic change that provides a moment of heightened energy and engagement" (Snoman, 2014). Here, the drop is spatialised: a sudden and dramatic shift occurs not only in tone but in the distribution of sounds around and above the listener. Høier (2020) similarly observes that "the dynamic panning of sounds – especially sound effects – can further heighten the 'sonic velocity'" and present desirable qualities, highlighting how surround channels and immersive audio tools extend the possibilities of spatial manipulation.

The transition from automated footsteps with height-channel rain to the static panning of percussion around the ear-level plane creates a striking contrast. This shift delivers both a tonal change, from the calm ambience of environmental sounds to the energetic texture of percussion, and a physical reorientation, as instruments now surround the listener on all sides. These elements combine to reshape the perceived size and acoustics of the listening space in ways that stereo formats cannot achieve. As Nosenko (2024) describes, "spatial placement of musical elements and acoustic modelling serve to increase or decrease the apparent size and distance to performers and sounds", demonstrating the creative and expressive potential of immersive tools.

Comparable strategies have recently appeared in film soundtracks. For example, Hans Zimmer's *F1* (2025) places a bass synthesiser more prominently in the side channels than in the front, creating a subtle sense of wraparound immersion. Immersio Overture develops this technique further, intentionally distributing percussion across the side and rear speakers. Whereas F1 employs immersive mixing to simulate a singular, enclosed performance environment, Immersio Overture transforms the listener's spatial perspective throughout the piece, continually altering the perceived position, room, and ensemble in motion.

2.2 VERTICAL AXIS SEPARATION

Creating an effect without conveying aural distraction is an important consideration. Part of the mixing is “to blend the auditory environment of a show so that listeners can accurately process the location of sounds without distraction or confusion" (Keyes, 2021). Within *Immersio Overture*, sounds with important or lead roles are positioned closer to the front of the mix, unless used in a supporting or effect-based role. Alternatively, positioning sounds left/right and forward/back allows for more dramatic effects. As Ziemer (2019) describes, “music lives and unfolds its effects with the room in which it resounds”, showing the philosophy of the space in a spatial environment primarily containing the “room” replacements and “effects” within. The positioning of sounds at 00:57 combines percussion instruments placed at ear level around the listener with lead brass and synthesiser parts playing in the height channels above. As shown in Figure 1, important harmonic and melodic elements are placed towards the front of the space, whereas supporting roles are placed further back.

Although distinct sonic layers are separated, this technique is not distracting; rather, it provides a powerful, creative, and technically coherent approach. By distributing instruments along the vertical plane according to their roles, it achieves clarity that, in conventional stereo or surround mixing, would require frequency or dynamic separation. This approach allows different layers to remain perceptually distinct without displacing the listener’s focal attention. At the same time, it creates an additional creative opportunity: immersion is deepened by affording sonic agency

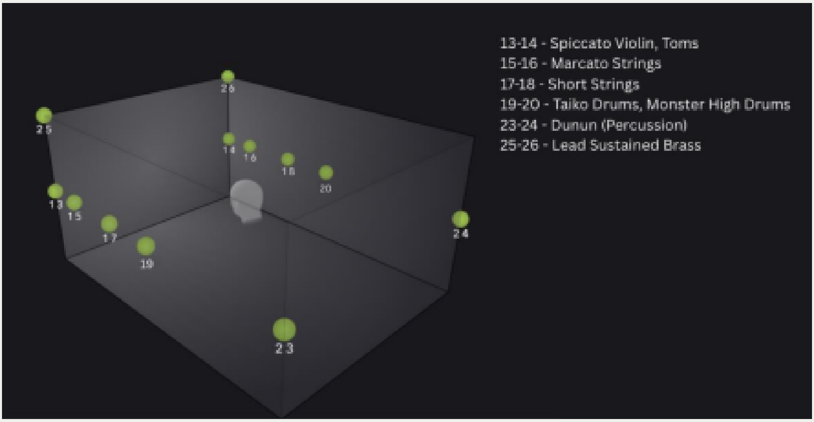


FIGURE 1. A VISUAL DIAGRAM OF THE CLIMAX AT 00:57 SHOWING THE PLACEMENT OF INSTRUMENTS WITHIN THE IMMERSIVE ENVIRONMENT

At the same time, it creates an additional creative opportunity: immersion is deepened by affording sonic agency across multiple horizontal positions, enriched by vertical differentiation. When the goal is to adjust density or to carve out ample space in the music without drawing focus away from key elements, vertical separation offers both technical and aesthetic advantages. Ziemer (2019) reflects on this in saying “Immersive audio compositions may use spatial sound to create the perception of expanding or contracting spaces, which can support narrative, emotional, or structural development within the piece". By providing controlled access to vertically panned sources, this technique enhances the detail of sonic positioning and opens a new dimension for composition. This is achieved not only through room recreation and emulation, but also by precisely affecting the direct sounds, thereby avoiding distractions for listeners.

2.3 NON-DISTRACTING HEIGHT ENERGY

At the beginning of the climax, at 00:57, harmonic instruments such as short and marcato strings, together with supporting synthesisers, are placed between the front and side planes at ear level. This placement is crucial for maintaining the listener’s focus, as important and foundational elements remain directed towards the audience.

At 1:24, new harmonic material in the form of continuous semiquaver figures drives propulsion and builds texture across the immersive plane. These elements are placed in the height speakers to prevent distraction from the primary activity in the front left and right channels. Their presence enhances immersion by expanding the sonic field above and around the listener. A central pulse initially occupies the height plane before diverging into two distinct sounds that animate the front and rear height zones. Figure 2 details the placement of instruments during the climax and demonstrates how spatial separation of roles creates clarity. At the conclusion of this section, each sound is automated into a separate corner of the height field, resulting in four separate textures articulating the same part simultaneously. This technique contrasts with the percussion, which remains anchored in the lowest horizontal plane. As these sounds are based on short-attack synthesizers, they were used to fill the space with non-distracting elements rather than to increase the density of texture.

In the piece Can You Hear the Music (Göransson, 2023), rhythmic synthesisers develop suspense and motion. These sounds are placed primarily at the front of the immersive playback system, although the sides, rears, and

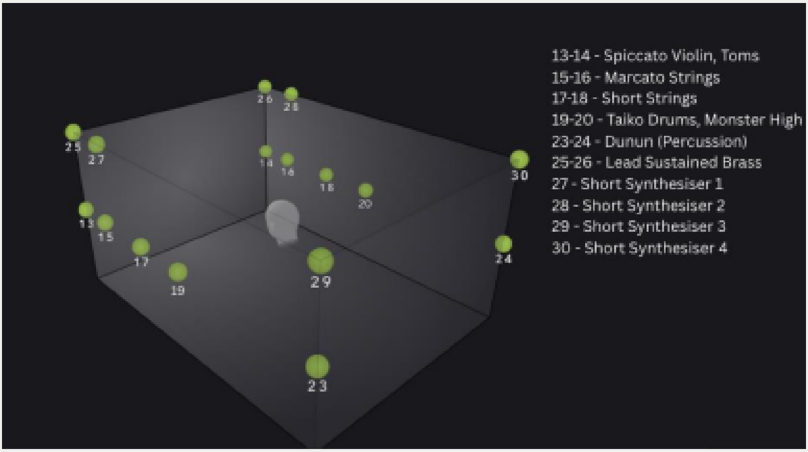


FIGURE 2. A VISUAL DIAGRAM OF THE CLIMAX FROM 1:24, SHOWING THE PLACEMENTS OF SYNTHESISERS

height channels also feature some of these elements. In Immersio Overture, this technique is extended by clearly defining instrumental parts and roles within the immersive system. The synthesisers are placed in the height channels to enhance clarity and avoid overlap with other elements. With subtle volume balancing, these sounds do not distract the listener but contribute additional energy. The purpose of this technique is to immerse the listener in new textures while providing non-intrusive rhythmic energy, adding depth without compromising focus.

Baxter (2022) observes that height channels “clearly contribute to the desired effect of total immersion”. Immersio Overture enables this immersion and further develops the textural variation by placing four different in the height channels to occupy specific corners of the system. As Baxter also notes, “irrelevant and distracting sounds are not typically elements of an effective sound design” (2022), an issue avoided in this case through the careful deployment of synthesiser timbres. Sounds with short attack times serve rhythmic functions; although more sustained than typical percussion, they provide both rhythmic and harmonic material. This enables individual parts and roles to remain clearly articulated, guiding the listener's attention toward the lead instruments.

2.4 MICROPHONE DISTANCE PANNING

In the final bars of *Immersio Overture*, at 01:50, a solo piano is heard. This part was deliberately shaped to emulate the impression of a distant piano, positioning the listener behind the performer. To achieve this effect, the Hans Zimmer Piano sample library from Spitfire Audio was used, which offers a range of microphone positions. Outriggers and surround microphones were selected to create a distant microphone signal and convey the impression that the audience's perspective was further away from the performers. Both signals were placed in the front left and right speakers and combined with additional reverberation routed to the surrounds and rear channels.

The reverb plugin Berlin Studio Professional was chosen for its flexible signal placement and the capacity to simulate traditional orchestral microphone arrays. In this configuration, the AB microphone signals were positioned to the sides, while the Surround signal was placed behind the listener, enhancing the impression of physical space and extending the distance from the piano sound. Figure 3 shows the microphone positions in Berlin Studio Professional relative to the piano, as well as available options such as the Decca tree, which was not employed in this project. The A-B pair and surround microphones offered greater perceived distance than the Decca tree, aligning with the compositional aim of making the piano sound as remote as possible.

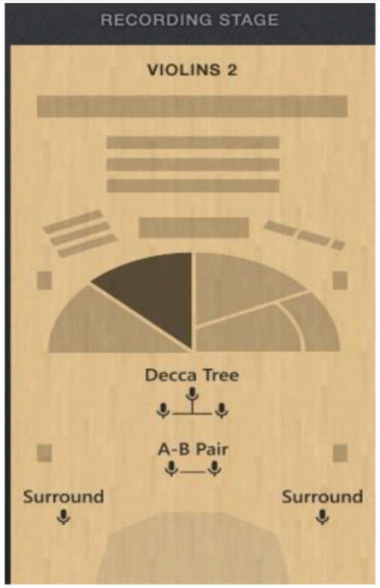


FIGURE 3. SCREENSHOT OF THE BERLIN STUDIO PROFESSIONAL PLUGIN BY SAMPLICITY, DISPLAYING A DIAGRAM OF A-B PAIR AND SURROUND SOUND MICROPHONE CONFIGURATIONS

The use of panning microphone signals has long been central to multi-channel formats such as 5.1 and 7.1.4. Haigh, Dunkerley, and Rogers (2020) describe how "left and right outriggers brought around the sides to a position about 30% of the way towards the rear speakers", illustrating one approach to distributing microphone signals in surround contexts. *Immersio Overture* builds on this tradition by experimenting with more extreme placements: for example, the piano's outriggers and surround microphones were panned to the front and then processed with convolution reverb to create a heightened sense of space.

These practices can be understood in relation to other immersive audio research. The project Exploring the Cinematic Hemisphere for Orchestra (ECHO) (Lindberg et al., 2025) demonstrates a range of microphone capture strategies specifically designed for immersive playback and panning. Lindberg's 2L Prism (2025), for instance, employs seven microphones at ear height and four height microphones, each spaced 100 cm apart, with the plan to route each signal to a corresponding speaker in a seven-channel ear-height array. In contrast, Willsher's P3H Anamorphic array pans the furthest signal back only 50%, limiting the use of the rear channels.

By comparison, Immersio Overture adopts a more expansive approach: reverb signals were placed across both the side and rear channels, drawing inspiration from projects such as 2L Prism while diverging from more conservative strategies like the P3H Anamorphic array. This decision was guided by the compositional objective of maximising the available speaker field to create a strong sense of space, distance, and clarity, while offering the listener as much unique spatial information as possible.

3 DISCUSSION

Research into immersive audio has frequently focused on the characteristics of acoustic spaces and the recreation of room environments, particularly in relation to how spatial design influences the listener's perception of immersion. For example, Ohgi, Miyazaki, Kim and Uhm examined Yamaha's REV system, which utilised 53 speakers to alter the "primary auditory characteristics of room acoustics" (Hiromu et al., 2023) in a performance space. Their study explored both the diffusion of sound through loudspeakers and the use of multiple microphone arrays. Four microphones were positioned close to each instrument of a string quartet (the In-line method), while twelve additional microphones were placed above the stage and in the audience area. This configuration enabled real-time manipulation of reverberation and sound-image control, allowing the acoustic environment to be adapted according to the demands of the performance.

Malyshev (2018) proposed a different perspective, recording full-band performances for 360-degree audiovisual presentation. His approach emphasised the importance of reproducing live environments, including their reflective properties: "We are reproducing a live environment that consists of reflections. Thus, it is important to recreate the room and its behaviour". The goal was to provide listeners with a stronger sense of presence within the VR environment through multichannel recording techniques, though the focus remained on accurate acoustic reproduction rather than extending or creatively manipulating the captured space.

Beyond these strategies of simulating realistic performance spaces, contemporary immersive loudspeaker systems also open possibilities for novel compositional and production approaches. In particular, the relationship between vertical and horizontal panning of direct instruments offers music producers new possibilities for shaping emotional impact and supporting the expressive goals of a composition, approaches that move beyond acoustic replication into explicitly creative spatial design.

Research on feature film music in 5.1 format shares common elements with hybrid orchestral practices. Holman (2002) distinguishes between a direct/ambient approach, where sounds “come from the front, and the surrounds are used for ambience”, and complete surround, which is “akin to placing the listener on stage or inside the world of the film”. While both approaches remain influential, they are primarily concerned with realism and cinematic convention. They do not typically account for hybrid or non-realistic strategies. In Immersio Overture, spatial placements are deliberately shifted: at times emulating a concert-like stage arrangement, and at others adopting more imaginative configurations, such as during the climax. These strategies suggest that immersive audio can move fluidly between representational and non representational modes, offering aesthetic possibilities that extend beyond filmic precedent.

Direct sound placement not only expands creative options but also delivers technical advantages. The techniques developed for Immersio Overture highlight how immersive audio can produce a clearer soundstage, increased dynamic variation, and improved frequency and dynamic separation, benefits less readily available in stereo. Rumsey (2016) notes that “upmixing two channel content for surround and immersive reproduction formats is therefore an attractive proposition if it can be made to deliver a convincing experience”, expressing a desire for improved experiences for the listener. Our work addresses this by positioning the listener within a drummer's circle, while assigning lead instruments to the height channels and situating an aggressive, sawtooth wave synthesiser at the centre of the listening space. This spatial arrangement generates horizontal variety while simultaneously maintaining vertical separation. In this sense, Immersio Overture demonstrates how immersive tools can be used not only to enhance realism, but to articulate musical structure in ways that stereo or conventional surround cannot.

Although research into height channel usage remains limited, some insights point to their creative potential. Baxter (2022) argues that “height speakers give greater resolution and creative possibilities in audio reproduction”, while Keränen and Hongisto (2010) emphasise the practical advantages of “allocating sounds like ambience, reverbs, or certain instrument stems [e.g., strings or synths] to the height channels, so composers and engineers can avoid overcrowding the horizontal plane”. Immersio Overture takes these perspectives further by using height channels not only for ambience or spill, but for clearly defined instrumental roles. This approach illustrates how the vertical dimension can be treated as an active compositional resource rather than a supplementary extension of the horizontal plane. In doing so, it contributes to ongoing discussions about how immersive audio may evolve from reproducing space to actively shaping musical experience.

4 CONCLUSION

Immersio Overture demonstrates how horizontal and vertical placements of direct sounds, when combined with distant microphone positioning and reverberation, can be used to construct spatial depth and shape musical perspectives. These techniques provide creative and technical benefits, such as enhanced layer separation and reduced reliance on intensive dynamic or frequency-based processing. Direct panning of instruments within the sound field allowed the composition to evoke the impression of a reconfigured ensemble in physical space. Additional strategies, including room reverb variation and the deployment of height channels, contributed to a more immersive environment than stereo, while horizontal distribution of contrasting elements facilitated clarity and prevented masking.

In response to the guiding question, “How effectively can selected immersive audio techniques be applied to hybrid orchestral music through composition?”, the four techniques considered (immersive impact, vertical axis separation, non-distractive height energy and microphone distance panning) each proved effective in ways that cannot be replicated in a 2.0 (stereo) environment. For instance, panning elements on the horizontal axis provided shifts in perspective, height placement enabled separation of parts and clear emphasis of lead elements; and microphone distance panning expanded the sense of depth. Collectively, these methods contributed to enhanced clarity, spatial interest, and new textural possibilities without introducing distraction.

It should be noted that many immersive works in hybrid orchestral genres, particularly those developed for streaming, tend to adopt film music conventions, prioritising room capture and realism over experimental spatial design. By contrast, Immersio Overture emphasised the compositional applications of immersive audio, moving beyond the replication of real-world acoustics towards creative manipulation of direct sounds. While a grounding sense of place was retained, the focus remained on exploring the potential of immersive audio as a compositional and production tool in its own right.

REFERENCES

Baxter, D. (2022). “Immersive Sound Production”. CRC Press. Available at: https://www.taylorfrancis.com/books/edit/10.4324/9781003052876/immersive-sound production-dennis-baxter [Accessed 25 Jul. 2025].

Bernschütz, B. (2016). “Microphone Arrays and Sound Field Decomposition for Dynamic Binaural Recording” (Doctoral dissertation, Technische Universität Berlin).

Brian, F.G.K., Poirier-Quinot, D. and Lyzwa, J.-M. (2021). La Vierge 2020: “Reconstructing a Virtual Concert Performance Through Historic Auralisation of Notre-Dame Cathedral”. IEEE Xplore.

Cunningham, S. (2018). “Proceedings of the Audio Mostly 2018 on Sound in Immersion and Emotion”. Audio Mostly.

Göransson, L. (2023). “Can You Hear the Music”. [Streaming] Apple Music. Haigh, C., Dunkerley, J. and Rogers, M. (2020). “Classical Recording”. CRC Press.

Hiromu, O., Hideo, M., Sungyoung, K. and Sihyun, U. (2023). “Transcending boundaries: Unleashing Musical Expression through Immersive Sound Image and Reverberation Control System”. [online] Journal of the Audio Engineering Society, pp.134–141. Available at: https://aes2.org/publications/elibrary-page/?id=22288 [Accessed 5 Sep. 2025].

Høier, S. (2020). “Surrounded by Ear Candy?” Nordicom Review, 35(s1), p.251–262.

Holman, T. (2002). “Sound for Film and Television”. [online] Boston: Focal Pr. Available at: https://www-taylorfrancis-com.ezproxy.newcastle.edu.au/reader/download/9bd7b10e829f-4bd2-9b0d-8e261801ece2/book/pdf?context=ubx [Accessed 7 Aug. 2025].

Keränen, J. and Hongisto, V. (2010). “Comparison of Simple Room Acoustic Models Used for Industrial Spaces”. Acta Acustica united with Acustica, 96(1), p.179–194.

Keyes, K. (2021). “Using Localization Cues in Immersive Mixing”. [Online] SoundGirls.org. Available at: https://soundgirls.org/using-localization-cues-in-immersive-mixing/ [Accessed 14 Apr. 2025].

Lee, H. (2019). “Psychoacoustics of 3D Sound Recording and Reproduction”. Zenodo, p.538– 551.

Lindberg, M., Wollage, N., Willsher, M., Lee, H., Ratcliff, S., Caruso, A. and Boynton, K. (2025). “ECHO Project”. [online] APL. Available at: https://apl-hud.com/echo/ [Accessed 27 Aug. 2025].

Malyshev, M. (2018). “Sound Production for 360 Videos in a Live Music Performance Case Study” . [Thesis] Available at: https://www.theseus.fi/handle/10024/150460 [Accessed 5 Sep. 2025].

Nosenko, O. (2024). “Immersive technologies in the Formation of Musical Audio Space”. Collection of scientific works ‘Notes on Art Criticism’, 24(1), p.23–32.

Rumsey, F. (2016). “Immersive Audio: Objects, Mixing, and Rendering”. Journal of the Audio Engineering Society, 64(7/8), p.584–588.

Snoman, R. (2014). “The Dance Music Manual: Tools, Toys, and Techniques”. New York, NY; London Focal Press.

Ziemer, T. (2019). “Psychoacoustic Music Sound Field Synthesis”. Springer. Zimmer, H. (2025). “F1”. [Streaming] Apple Music.

SECTION III - ADVANCES IN DIY HARDWARE STRATEGIES

Smart Musical Mats: a musical ubiquitous computing artifact to support basic education (Santos; Filippo; Pimentel);

Maudlin: Interface for Random Walks and Beyond - Eurorack Module for Ubiquitous Music (Harding);

Standalone Interactive and Generative Music with the Csound-FPGA Framework (Jagwani; Lazzarini).

SMART MUSICAL MATS:

A MUSICAL UBIQUITOUS COMPUTING ARTIFACT TO SUPPORT BASIC EDUCATION

THIAGO MARCONDES SANTOS[1] · **DENISE FILIPPO**[2] · **MARIANO PIMENTEL**[3]

THIAGO.MARCONDES@UNIRIOTEC.BR PIMENTEL@UNIRIOTEC.BR DFILIPPO@ESDI.UERJ.BR

ABSTRACT

The problem investigated in our research is the difficulty posed by technical barriers to sound making in the process of learning music, and this problem is located within a sociotechnical context (basic education). In order to investigate the solution to this problem, we have developed a ubiquitous computing artifact, the Smart Musical Mats (SMM). In this article, we present the developed artifact, the results of our research, developed using the Design Science Research approach, and also our thoughts on how this artifact fits within Ubimus.

1 INTRODUCTION

An important part of music education is the making of music by students. However, some musical instruments require difficult techniques to make sounds, and demand a high number of exercises and of practices with repetitive movements which, often, cause students

---

1,3. PPGI-Postgraduate Program in Informatics – Federal University of the State of Rio de Janeiro (UNIRIO), Pasteur Av. 458, Ground Floor, Urca, 22290-240, Rio de Janeiro, RJ, Brazil.
2. Superior School of Industrial Design (ESDI) – University of the State of Rio de Janeiro (Uerj), Evaristo da Veiga St, 95, Centro, 20031-040, Rio de Janeiro, RJ, Brazil

to give up on trying to play them. Therefore, one can observe a problem that will need to be faced while implementing music education in Brazilian basic education. How can we support sound making by students in order to decrease technical difficulties and make the process of learning music more pleasant and less dependent on instrument technique? This issue is even more relevant in the first years of basic education, when children have not yet developed aspects of laterality and motor coordination which are essential for playing certain musical instruments. In order to solve this problem, we developed a ubiquitous computing artifact, the Smart Musical Mats, which was also investigated and evaluated in the work of Santos (2015), who is one of the authors of this article and who was also the researcher who conducted the case study presented here.

The technological changes that computers have brought to everyday life are multiplying in several application fields. This multiplicity of options is accentuated when computers also physically spread through the environment. Weiser (1991) foresaw this moment and coined the term Ubiquitous Computing. He understood that computer components becoming smaller would lead to a revolution in our way of perceiving and interacting with computers in our lives: each person would have not only one, but several computing devices. His vision proved right, because today we are living a time when we have not only smartphones, smart tablets and smart game consoles, but also smart personals assistants such as Google Home and Alexa, smartwatches, smart vacuum cleaners, smart locks and smart cars. This new way of interacting with computers brings many possible applications in different areas such as music, education and, consequently, music education.

For over a decade, a research field called Ubimus (Keller et al., 2019) has been investigating different types of concepts, constructs and artifacts linked to the combination of ubiquitous computing technologies and the field of music.

Section 2 deals with what the authors of this article consider music education and its goals. Section 3 deals with Ubiquitous Computing and the developments of its relationship with music, including examples of Ubimus' lines of research and important concepts that have emerged from these researches. Section 4 presents an artifact that was designed and investigated in context in order to support classroom music education in basic education, the Smart Musical Mats, and Section 5 presents the methodological and epistemological aspects of this research. Section 6 presents the results of the investigation and the evaluation of the artifact, and Section 7 contains a few thoughts on the possible relationships between the artifact and Ubimus. Section 8 presents the conclusions of the article, limitations and future works.

2 MUSIC EDUCATION AND THE THEORETICAL FRAMEWORK OF THIS RESEARCH

In order to better understand the goal of music education in basic education, we need to define our understanding of what is music education and how is the relationship of students with music throughout their learning process. Thus, we have established a premise and three conjectures which, derived from our theoretical framework, have guided the design of our artifact.

Our premise consists in the consideration that music is a type of discourse, and that music education's objective is to make students literate in musical discourse. The three theoretical conjectures that were brought from our theoretical framework were based in the relationship between the technical demands of sound production and the musical learning process, in the social aspect of education and in the use of the body and its movements in developing music education. Below, we present the premise and theoretical conjectures that guided the development of the artifact presented in this article.

2.1.1 MUSIC AS DISCOURSE

Music is a sequence of sounds in which different sounds cause listeners to experience different sensations, emotions and thoughts. Thus, we can understand music as a manifestation that belongs both to the physical universe of sounds and to the emotional and cognitive universe. Some authors believe that music has a communicational side and that it is similar to language in some aspects.

2.1.2 SWANWICK'S MUSICAL DISCOURSE

Swanwick (1979) presents the notion of musical discourse. Thus, for him, music is composed of sound elements that are combined according to a common value and a common intention perceived by the participants, who are within a creative process that must be fluid and contain meaning. He presents a few important concepts regarding the most basic elements of musical discourse, which we describe below:

- **Materials** are any sounds that can be perceived, but are not intentional and do not have a relational character, such as, for example, the sound of rain, the sound of a glass breaking, a musical note (of a given pitch) played randomly (such as when a musician tunes his or her instrument).

- **Elements** are any sound materials which, due to the intention of the player, acquire a relational character and become an element of the musical discourse.

In order to reach the level of musical element, the material must have three different conditions (Swanwick, 1979):

- **Selection** – In order to build a musical object, one must choose certain sounds and discard others. Not all available sounds are selected, and there are situations in which certain sounds can be discarded and others, repeated.

- **Relation** – The chosen sounds must have a given relationship with one another in time.

- **Intention** – The composer/performer must have the intention of making music.

The way in which sound making occurs in a musical instrument has a direct impact on the student's ability to make sound materials and in their transformation in musical elements that can be used to build musical discourse.

Therefore, the difficulty faced by the student in making sounds using a given musical instrument will influence their way of expressing themselves and of participating in building musical discourse.

2.2 DECREASING TECHNICAL BARRIERS TO THE BUILDING OF MUSICAL DISCOURSE

Educators such as Orff (Frazee & Kreuter, 1997) tried to make sound making easier by offering students instruments such as the metallophone and the xylophone, since they allow sound making without a great deal of technical effort. The xylophone is also built with bars that can be removed, and it can be configured so as to avoid certain mistakes on the part of the student on a given exercise: for example, with the removal of certain bars, you can leave only the specific subset of notes with which they will work in that class. By doing this, you avoid technical performance issues such as exceeding notes that can lead the student to make a mistake. Thus, we consider our first conjecture: decreasing technical barriers in sound making supports musical discourse literacy. The research problem is related to this first conjecture.

2.3 THE SOCIAL ASPECT OF EDUCATION

To Vygotsky (1989), social interaction is determinant in the learning process. The subjects and the object of knowledge create different types of relationships that include interactions between a subject and the object being studied, among other subjects and the same object and, most importantly, among one subject and the other subjects that interact with that object. The social and collaborative relationships among subjects potentialize and intensify the discoveries regarding the objects being studied. Based in his studies and researches on social interactions and learning, Vygotsky then elaborates the concept of Zone of Proximal Development (ZPD) as being:

> (...) the distance between the real development level "of the child", determined by independent problem solving, and the potential development level determined by problem solving under the guidance of adults or in collaboration with more knowledgeable others. (Vygostsky, 1989, p. 86).

The learning process, then, is a cultural consequence, derived from the contact among the subject and other subjects who also manipulate and experiment with the objects of study. Based on this characteristic of education formulated by Vygotsky, we have arrived at our second conjecture: *collaboration supports musical discourse literacy*.

2.4 THE RELATIONSHIP BETWEEN THE BODY AND THE MUSIC LEARNING PROCESS

Authors such as Dalcroze (1921) and Orff (Frazee & Kreuter, 1997) also defended the use of the body and the use of movement through gestures as mediators of the music learning process, supporting the understanding and the assimilation of musical discourse. Learning activities that privilege both intellectual and physical aspects have a greater impact in the perception and in the learning process of individuals involved. Supported by these authors, our third conjecture is: the use of the body and its movements supports musical discourse literacy.

After presenting our understanding of the relationship between people and musical discourse, in the following section we will discuss the possibilities of computer musical artifacts that can be built using the technological advances in the field of computing that occurred in the last few years.

3 MUSIC AND UBIQUITOUS COMPUTING

The relationship between music and ubiquitous computing is relatively recent and has been investigated for over a decade. "Ubiquitous music is, in practice, music (or musical activities) supported by Ubiquitous Computing concepts and technologies" (Pimenta et al., 2015). Ubimus' research is characterized by different fields of knowledge and investigation, including "sound and music computing, human-computer interfaces, studies about creativity and music education, with a strong social and communitarian component in its practices" (Keller et al, 2019). These researches deal with the participation in artistic musical contexts (Keller & Capasso, 2006) and in educational contexts (Lima, Keller & Flores, 2018; Keller & de Lima, 2018; Miletto et al., 2011) by both laypeople and subjects who are already familiar with music technique, such as professional musicians or experts.

Important concepts are present in Ubimus' research, such as everyday musical creativity and ecological-cognitive creative practices. Based on the concepts of little-c creativity

(Richards, 2007), Keller et al. (2013) state that everyday musical creativity is a specific form of everyday creativity that is “derived from creative sound processes and products that occur in the everyday life of musicians and laypeople outside of environments specifically designed for the making of music.”. Ecological-cognitive creative practices are the result of the interaction between subjects, with their cognitive processes, and the material resources present in their environment (Keller & Lazzarini, 2017).

Keller and Barreiro (2018) also call attention to the existence of two predominant forces of attraction in Ubimus’ research: studies about everyday (musical) creativity and new ways to make music.

The first force of attraction is everyday creative practices, in which there are issues related to user profiles and the changes that these different profiles cause on the design of the interactions present in systems developed by Ubimus. Users who can be considered laypeople regarding musical knowledge and ability create demands that are different from those created by experienced users. Therefore, different types of system can be developed to meet the demands of specific user profiles. However, problems can arise when two different student profiles are mixed in the same musical activity, a situation that still needs to be investigated and evaluated. A possible solution to this problem is the creation of knowledge transfer strategies among laypeople and experts through gesture cues made by body movements (Keller, Aliel & Siva, 2018). Another possible solution to the same problem is the implicit knowledge transfer among participants supported by verbal language, in avoiding the use of specific musical technical terms. This latter approach can also lead to the inclusion of semantic cues in Ubimus' ecosystems, an example of which is the tool Playsound.space (Stolfi et al, 2018).

The second force of attraction is new ways of making music, and this force has three different aspects:

1. The change in the limits of the interactions among composers, audience and performers during musical performance.

2. The potential impact of Ubimus in human development and in the wellbeing of its users.

3. The influence of ubiquitous computing techniques in the development of Ubimus.

Regarding the third aspect mentioned previously, the advances in ubiquitous computing allowed the emergence of different technologies such as 3D printing, laser cutting and the use of microcontrollers and processors that can, due to their small size, be placed in different locations and objects. Maker culture and the concept of DIY became a part of possible Ubimus practices.

Brown & Ferguson (2024) bring up good questions about the use of digital fabrication technologies and the use of ubiquitous computing to build sound instruments. In the case of their instrument, Analogue Revolutions, the authors built it using integrated circuits (ICs) and passive components, stating "*there are no programmable (PIC) chips and thus no computer is involved.*". In the case of another instrument, Sonic Frisbee, the authors' objective was to use it in workshops and other pedagogical situations that include both sound-musical practice and the maker culture practices of becoming familiarized with welding processes and of making sounds with electronic components. Their other creation, Quadra, is a kind of *Groovebox1* more appropriate for being used by composers than performers. The authors used laser cutting machines to build the printed circuit boards (PCB) that were used in these sound instruments.

Merendino (2004) built the Dispositivo Cinetico Midi (DCM), a digital sound instrument that has buttons that can be finger-activated, but added motion sensors, using a gyroscope, that allow a larger number of gestures to influence in sound production. Merendino (2024) also presented an instrument named Sonic Cubes in which cubes made out of fabric have an esp32 microcontroller inside. Each cube is wrapped in a piece of fabric that has conductive thread running through it and acts like a user touch sensor. When they are moved, the cubes send data captured from the gyroscope to a computer, in the same environment via Bluetooth, and produce sounds.

4 SMART MUSICAL MATS

The developed artifact, named Smart Musical Mats (Santos, 2015), is composed by a set which contains something that is a part of everyday school life, the EVA foam mats, a hardware and a software. We gave one object – the mats – a certain degree of intelligence, connecting it to different computing elements and sensors, all of which need to be present in the classroom.

When we consider the taxonomy proposed by López and his collaborators (2009), the artifact Smart Musical Mats presents abilities 1 (it has a unique identity), 2 (it perceives changes in the environment, i.e. electric circuits closing), 3 (it acts on the environment, making sounds) and 4 (it decides which sounds to make depending on the perception of which electric contact was closed).

4.1 THE OBJECT OR MAT

Squares of EVA foam shaped like pieces of a puzzle are an everyday object on different segments of education. From day-care to secondary school, we can observe these objects being used for classroom decoration, to protect children from falls and in games and playful activities in general. These EVA foam squares are usually sold in sets, for example of four, six, ten or twelve, and are often used to make carpets. Thus, in this article, we have considered each square as a mat and called our artifact Mats, plural, because it needs at least two squares to work. On Smart Musical Mats, each mat is connected to the hardware. EVA foam mats are inert objects. The hardware and the software connected to them is what makes them a smart object with which subjects can interact.

4.2 THE HARDWARE

The hardware built during this research is composed by the following parts: a computer, electric wires, aluminium foil sheets, a Makey Makey (2020) board, which will provide the system input, and loudspeakers, which will be the system output.

The Makey Makey board has an electric configuration with eighteen signal input pins and also ground pins. This board has an embedded ATmega32U4 microprocessor, the same one used on the Arduino Leonardo, which allows it to emulate key presses and mouse movements using the HID (Human Interface Device) protocol. These mouse movement or key press commands are turned on when the electric circuits are closed by the presence of any minimally conducting material. Thus, each pair of input and ground pins works as an on/off switch, which begins or interrupts a HID command addressed to the computer to which the Makey Makey is connected.

In the case of the Smart Musical Mats, we used the human body as the conducting material that closes the contact between an input and a ground pin. More specifically, we used two people, each connected to a wire. When they touch each other, they close the circuit associated with them. Working in pairs, each person, barefoot or wearing socks, stands on the aluminium foil sheet over a mat (Figure 1). The sheet connects to a wire that is extended to an input pin on the Makey Makey board.

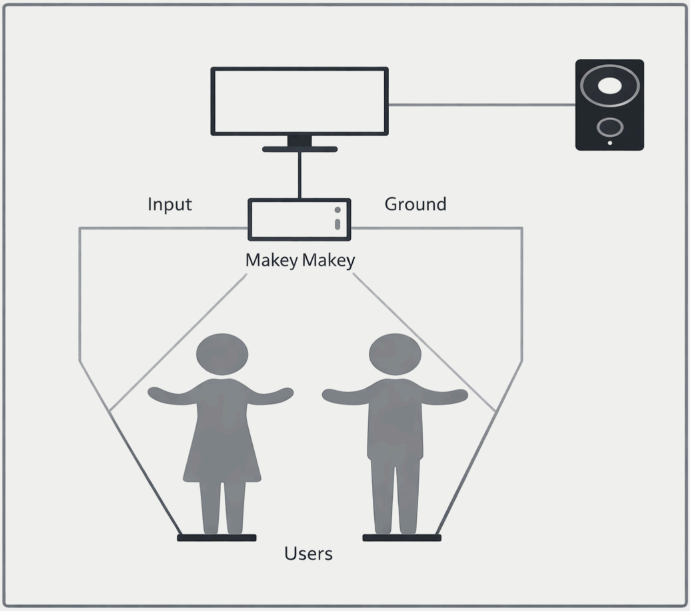


FIGURE 1. STUDENTS POSITIONED OVER THE MATS

4.3 THE SOFTWARE

Soundplant (2020) is a proprietary software, but it has a free version which was used in this research. This software addresses different audio samples (in formats such as wav, aiff, mp3) to the computer keys and, when one key is pressed, a specific sound is triggered. The Makey Makey board, then, can activate the Soundplant software, and the audio samples chosen by the users are triggered in the computer and heard through loudspeakers.

5 METHODOLOGICAL AND EPISTEMOLOGICAL ASPECTS OF THE RESEARCH

The production of knowledge about the development of artifacts is the objective of the Sciences of the Artificial (Simon, 1969), which must be studied in a different manner from the Natural Sciences and the Social Sciences. Since this work investigates an artifact that does not exist in nature, but was created by man, we understand that this research is part of the Sciences of the Artificial, as detailed in this section.

5.1 METHODOLOGY-EPISTEMOLOGY

The technical knowledge generated during the making of an artifact is different from the scientific knowledge generated by a research (Pimentel et al, 2020). In order to do a research that generates scientific knowledge during the development of our artifact, we used the epistemological-methodological approach Design Science Research – DSR (Hevner et al, 2010) (Figure 2). This approach, which is based on the Sciences of the Artificial, has the double objective of: (1) developing an artifact to solve a practical problem within a specific context and (2) generate new technical and scientific knowledge (Dresch et al., 2015). In a research that uses the DSR approach, the development of the artifact supports the advance of the theory on user behaviour, which, in its turn, supports the development of the artifact. This occurs on an iterative process in which both are mutually refined. In DSR, one uses the abductive thought typical of a projective work, in which alternatives are identified and solutions are selected, differently from what occurs in inductive, deductive or hypothetical-deductive thought.

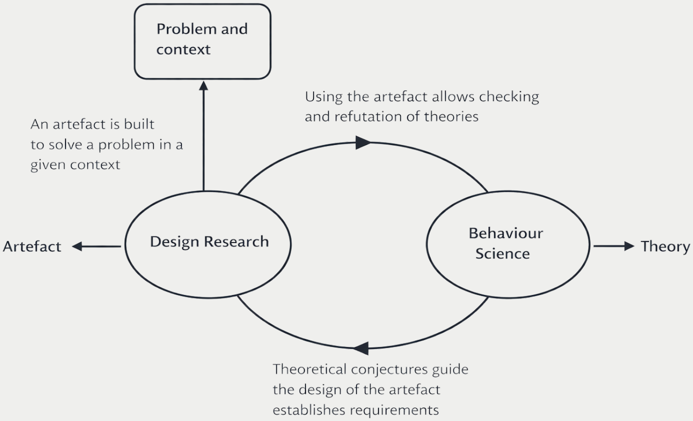


FIGURE 2. DSR ELEMENTS

In this research, we developed a didactic musical object (Design Research) denominated Smart Musical Mats (Artifact) in order to solve the difficulties imposed by the technical barriers to sound making in the process of learning music (Problem) within the sociotechnical environment of basic education (Context). At the same time, during the development of the artifact, we investigated theoretical conjectures about the process of teaching music based on the uses and the relationships of participants with this artifact (Behaviour Research) within this new context in which we utilize an artifact supported by ubiquitous computing technology. Thus, this research deals with a real problem, an artifact that aims to solve this problem, the development of this artifact, and the behaviour of those who use it.

5.2 DATA PRODUCTION

We used the Case Study research method following a qualitative approach with an interpretationist epistemological position (Yin, 2015; Pimentel et al., 2017) typically used in the Social Sciences: its objective is deriving values and knowledge about people's behaviour in a given context through the interaction of the researcher with the research participants. The case study was done in two year 7 classes in municipal schools located in the city of Rio de Janeiro, in Brazil, with a total of 75 students.

This case study began with a conversation with the teacher of each school, which allowed the author to watch one the classes. After this initial contact, four educational activities were developed to support the investigation of acceptance and use of the artifact, and the investigation of our theoretical conjectures on how people learn music. The same four activities were performed in both schools as pedagogical workshops that took place during music classes. These activities, performed throughout a period of approximately three months during the year 2014, were led by one of the authors of this research under the supervision and with the support of the music teachers of both classes.

The technique of data production was performed at the end of the four activities through a questionnaire and a focus group for the students, and through interviews with the teachers. Another technique used was the direct observation by one of the authors of the activity while it was happening and, later, through a video recording of it.

The questionnaire was composed by close-ended questions and open-ended questions that were related to the student profile and their previous music experience. The focus group questions meant to find what the students thought of the equipment, whether they had any difficulty using it, what they learned from the activity and whether they would like to have another class that included the artifact.

The questions of the interviews with the teachers were semi structured, with defined topics that were discussed freely. The topics regarded the teachers' views on music education within basic education with the arrival of Law no. 11.796/ 2008, the use that their students made of the artifact and whether or not they would like to use the artifact in their classes.

We used the Method of Clarification of the Subjacent Discourse (Nicolaci-da Costa, 2007) to analyse and interpret the data obtained from the focus group and from the interviews with the teachers.

This research was approved by the ethics committee of the Federal University of the State of Rio de Janeiro under the report number 893.135 given by the Report Office on 11/26/2014.

5.3 PEDAGOGICAL WORKSHOP ACTIVITIES

We organized four musical activities to be performed during the pedagogical workshop done for the case study. The activities were designed to support different aspects of music education mentioned in the National Curriculum Parameters (PCNs, 1998) presented by the Brazilian Ministry of Education.

5.3.1 RHYTHMIC ACTIVITY

The rhythmic activity (Figure 3) involves the pulse and its divisions in the composition of rhythms, and the appreciation of characteristic sounds (timbres) of different instruments used in the rhythmic accompaniment made by a drum kit (bass drum, snare and hi-hat). In this rhythmic activity, we used three pairs of mats with one drum sound for each pair (bass drum, snare and hi-hat). The students took turns so that all could try using the artifact and producing the required sounds (Execution). Three rhythmic lines were presented to the students using handclaps only, so that later they could be played using the artifact (Appreciation). The three rhythmic lines were built (Composition) in order to accompany songs with which the teachers of each school were working.

FIGURE 3. STUDENTS DURING THE RHYTHMIC ACTIVITY

5.3.2 MELODIC ACTIVITY

The melodic activity (Figure 4) involves listening to the sound of different musical instruments (bass and violin) and perceiving the major scale accompanied by the performance of cyclical melodies (ostinatos) among groups of students. In this activity, we distributed ten pairs of mats, with each mat representing one of ten notes played by a violin (C3 to E4). First, we asked students to play the sequence of notes from the lowest to the highest pitch. Then, two groups, each with five pairs of mats, were given the task of creating melodic sequences (Composition) so that the class could appreciate their compositions (Appreciation). The sequences were performed following the pulse played by the teacher and the claps, and later the students were given more rhythmic freedom (Execution).

FIGURE 4. STUDENTS DURING THE MELODIC ACTIVITY

5.3.3 HARMONIC ACTIVITY

The harmonic activity involves the perception of the difference between single notes and chords, the concept of a block chord and an arpeggio and the harmonic accompaniment of melodies (Appreciation). We used the same structure of ten pairs of mats as in the melodic activity. Each group of five pairs contains the four notes of a tetrad (a four-note chord), with one note per pair. The last pair of mats has the sound of the chord with all notes played simultaneously (block chord). The notes had to be played as an arpeggio by the students using the first four mats and, afterwards, the fifth pair had to play the full block chord, so that they could perceive the difference between the block chord and the arpeggio (Execution). In this activity, we used piano sounds. In the beginning, the rhythm was fixed, following the same pulse, but later the students were given more freedom. Afterwards, we did a variation of the activity and changed the sounds for each pair of mats. One group of mats contained three chords (C major triad, F major triad and G major triad) distributed over three pairs. Two of the five pairs were not used in this variation of the activity. The other group of five pairs was configured with five vibraphone sounds containing the notes of the C pentatonic scale (C3, D3, E3, G3, A3). The teacher asked the students of each group to create their sequences freely using the notes, chords and rhythms (Composition), so that everyone in the class could hear a melody accompanied by chords.

5.3.4 ORCHESTRA ACTIVITY

This activity's objective was joining the rhythmic, melodic and harmonic elements involved in the other activities and allowing the students to accompany the song being taught in the class by the teacher with a rhythmic, melodic and harmonic basis. The focus of this activity was performance and rhythmic activity, not composition, as proposed in the melodic and harmonic activities. Another objective was developing the melodic memory of students and showing the difference between accompanying and singing a song.

In this activity, we distributed sounds of different instruments among the pairs of mats (Figure 5) so that all the rhythmic, melodic and harmonic elements involved in previous activities were combined. One group had the drum sounds, another group had the bass sounds and another group had the chords played by a piano. The class was split in two larger groups and, while one sang the song, the other accompanied it with the sounds of drums, bass and chords. Afterwards, the two larger groups switched: those that were singing began accompanying, and those that were accompanying began singing.

FIGURE 5. STUDENTS DURING THE ORCHESTRA ACTIVITY

5.4 EVALUATION OF THE ARTIFACT

We did three evaluations with the artifact: acceptance of the artifact by users, whether or not the artifact solved the problem of technical barriers to sound making with an instrument, and an evaluation of the conjectures about music education brought from our theoretical framework.

5.4.1 ACCEPTANCE OF THE ARTIFACT BY USERS

Venkatesh and his collaborators (2003) investigated different models of technology acceptance and synthesized eight of these models: the Theory of Reasoned Action (TRA), the Technology Acceptance Model (TAM), the Theory of Planned Behavior (TPB), a model combining TAM and TPB, the Innovation Diffusion Theory, the Social Cognitive Theory, the Motivational Model (MM) and the Model of PC utilization (MPCU). As a result of the synthesis of these models, the authors presented the Unified Theory of Acceptance and Use of Technology (UTAUT). The UTAUT model was developed in order to investigate technology acceptance and usage in a business context, with four constructs being considered: *Performance Expectancy, Effort Expectancy, Social Influence* and *Facilitating Conditions*. The three first constructs influence the *Behavioural Intention*, another variable that influences *Use Behaviour*. The *Facilitating Conditions* construct directly influences *Use Behaviour*.

Later, Venkatesh and his collaborators (2012) proposed the UTAUT2, an extension of the UTAUT model, adding three other constructs that influence technology acceptance and use by consumers: Hedonic Motivation, Price Value and Habit.

In order to evaluate the acceptance of the artifact, we used the extended technology acceptance model UTAUT2 and considered that the artifact Smart Musical Mats should meet the following constructs: having good performance (allowing the adequate exploration of musical discourse), having low use effort (being easy to use), promoting hedonic motivation (being fun to use), leading to use behaviour (making students want to use it in other classes) and having low cost. We did not consider Habit as a relevant construct for the artifact proposed here, because this specific artifact had never been built before. Regarding the Social Influence construct, we consider that society, parents and students already, to a certain extent, pressure schools to use computing technologies in education and, thus, the fact that we are designing an ubiquitous computing artifact would already lead to use behaviour; therefore, this construct was not investigated in the study done with the artifact, even though it has a positive influence over its adoption. The model also includes the construct Facilitating Conditions, which deals with how easy to build and operate is the artifact, and this construct was evaluated in the interviews with the teachers. These constructs guided the questionnaire's questions and the questions asked during the focus group and the interviews with the teachers.

5.4.2 EVALUATION OF THEORETICAL CONJECTURES AND RESEARCH PROBLEM SOLUTION

The evaluation of theoretical conjectures was done through the direct observation and the participation of the researcher, who is one of the authors of this article, during the musical activities performed in the workshops and later through the examination of the videos recorded during the activities. The evaluation was also based on the data extracted from the focus group composed by students and the interviews done with the classes' music teachers. The objective of these evaluations was answering if our three theoretical conjectures were valid:

- *Conjecture 1 – decreasing technical barriers in sound making supports musical discourse literacy.*

- *Conjecture 2 – collaboration supports musical discourse literacy.*

- *Conjecture 3 – the use of the body and its movements supports musical discourse literacy.*

6 RESULTS

Here, we will present the reports from students considering our three theoretical conjectures and the constructs of the UTAUT2 technology acceptance model. The analysis of all this information will allow us to evaluate whether our artifact solved the research problem. At the end of this section, the results of the questionnaire given to the students will also be presented.

6.1 DECREASING TECHNICAL BARRIERS SUPPORTS MUSICAL LITERACY

Students thought the artifact made sound making easy and that they could pay attention to the musical discourse: *"I just had to wait for the right moment to clap. Sometimes it was hard, but then I started to get it"*. In the interview with the teacher of this class, the researcher asked whether he thought that it was easy for students to use the artifact, and he considered that the difficulty of making music with the artifact is the same as with any other instrument: *"I think they had the usual difficulties a person has when they're learning a new instrument. They have that initial block, but then it starts to develop more easily"*. However, the students had a different opinion and considered that the artifact was easier to play than traditional musical instruments: *"I liked it because it's a technological instrument and it's easier to play. Before you learn how to play an instrument, it's kind of hard and you have to practice. I liked participating in an experiment like this one"*; *"I thought it was cool, because I thought it was easier to play and easier to learn. It's easier to pay attention* *because with the others, with traditional instruments, it takes too long to learn”.* Since most of the students have traditional musical instruments at home (according to one of the results of the questionnaire), the opinions of these students are based on experience, which indicates that the artifact really allowed them to make sounds more easily than traditional instruments. This indicates that the artifact contributes to the solution of our research problem, although we cannot unequivocally state that it has been solved. Future works can investigate the use of this artifact in other contexts, giving a better understanding of the solution of the problem that originated this research.

6.2 COLLABORATION SUPPORTS MUSICAL DISCOURSE LITERACY

A few of the students highlighted the interaction established with their classmates: *“I liked it a lot because it’s very… I interacted a lot.”; “I thought it was really interesting (…) Especially because we work as a group and the sound is much better. It’s not just one person working alone. Instead of learning stuff by heart, we get the rhythm, so it’s more interesting”*. These statements identify the social aspect of building musical discourse.

In the interview with the teacher of one the classes, he emphasized the importance of collaboration among students in the learning process: *“I saw students collaborating a lot! You can use more than one child to make a sound. (…) The coolest thing is that a student depends on another to make the sound. If one of them has more ability and other has less, it can make things a little difficult in the beginning, but then there’s also the issue of one helping the other develop.”*

6.3. THE USE OF THE BODY SUPPORTS MUSICAL LITERACY

The use of the body was highly exercised during the activities. Students used movements to express themselves and a lot of them identified these movements with the learning process: “I learned the rhythms through the body movements”; “I learned motor coordination, how to keep time, before, I didn’t play at the right time. Even the teacher complained that we played too fast, but then we got used to it and got better”; I learned motor coordination and the time for each note [claps]. Synchronizing is waiting for the right moment to clap”.

The researcher also observed a situation in which two students who were not using the mats danced to the music, and another in which a student, after the activity, tried to make sounds sliding his body through the floor until he reached the mats in different ways. The way the mats were arranged in the classroom environment allowed students to change position and move freely, interacting with the artifact and also with their classmates during the learning process. A few students used movement strategies to fill pauses between the sounds played in their activities, indicating that their bodies and their movements were used to build the musical discourse in an active and also in an authorial manner.

6.4 ACCEPTANCE OF THE ARTIFACT

Regarding Performance Expectancy, the activities done with the artifact allowed a large variety of learning experiences among students. A few students highlighted what they had learned during the activities: *“I learned the chords, the times* *we use. I don't know… I learned how to play the instruments without having to use the real ones".* Regarding *use effort,* students stated that *"In the beginning it was hard to follow the rhythm, but then I got it right"; "In the beginning my coordination wasn't good and I played too fast. But then it got better".* Even if they had initial difficulties, students didn't feel intimidated and wanted to explore the artifact. *"It was a little difficult, but it wasn't difficult like… oh… that was really hard".*

Regarding *Hedonic Motivation*, students stated: *"To me that's a new instrument and I liked it a lot, a lot, a lot!!"; "At first, I thought it wasn't going to work. Me, making music with my hands!! [laughs] But then I saw that it was really fun and cool and that we learn. In the beginning, I was afraid to get shocked".*

The only construct that did not get good results in the investigations was Facilitating Conditions. Our artifact is a prototype and needs improvements in some aspects of its setup process so that anyone who is not an expert in informatics can set it up to perform activities quickly and easily. In the interviews, the teachers said that they liked the activities and that they liked seeing the artifact being used by students, but that they did not feel able to set it up without help. Thus, future investigations could focus on other ways to set up the artifact in order to make it easier to use for teachers with basic computer knowledge.

Regarding low cost, all the materials involved in building the SMM are cheap and easy to find: electric wires, aluminium foil, duct tape, EVA foam mats, adhesive tape. A Makey Makey board costs less than £40,00 and Soundplant has a free version. A pair of mats ready to be connected to a

Makey Makey board costs approximately £4,00. These costs are perfectly feasible for most municipal schools in Rio de Janeiro and in several other medium-sized and large-sized cities in Brazil.

6.5 QUESTIONNAIRE DATA

We gave students a questionnaire to answer after the last workshop activity. The first part of the questionnaire had questions about the students' profiles and the second part had questions about the activities and the use of the artifact. While some students answered the questionnaire, others were simultaneously interviewed in the focus group, so that the data collecting process took less time and did not hinder the schools' schedule and the classes planned by other teachers. The questions in the second part of the questionnaire presented answer alternatives in a Likert scale with five values dealing with issues such as difficulty, joy and satisfaction regarding the activities and the use of the artifact.

On *School 1*, 75% of students had between 12 and 13 years of age, and 54% were male. On this school, 68% already had a musical instrument at home and 61% learned music exclusively at school. On *School 2*, 46% of the students had 13 years of age and the remaining students had 12 years of age. 46% were female. 69% of the students of this school did not have musical instruments at home and 77% learned music exclusively at school.

Regarding the activities with the use of the artifact on *School 1*, 60% of students thought they were easy and 10% thought they were difficult. On School 2, 85% of students thought they were easy and none of them thought they were difficult. Regarding joy in using the artifact in the activities, two questions were asked: whether it was pleasant and whether it was fun. On School 1, 71% of students thought it was pleasant and 10% thought it was unpleasant. 71% thought it was fun and 10% thought it was boring. On School 2, 69% of students thought it was pleasant and 8% thought it was unpleasant, while 69% thought it was fun and 15% thought it was boring. When asked if they intended to use the artifact again, on School 1 75% answered yes and 25% answered no. On School 2, 85% answered yes and 8% answered no. These answers support the analyses about the UTAUT2 constructs Performance Expectancy, Hedonic Motivation and Use Behaviour, indicating that the experience of using the artifact was easy to perform, pleasant and made most participants want to use the artifact again.

These results, when evaluated along with the data gathered by the focus group and the interviews with the teachers, indicate that the artifact has acceptance potential among students and teachers.

7 THOUGHTS ON THE SMART MUSICAL MATS AND UBIMUS

After presenting the artifact and the research with which it was possible to evaluate and investigate its use by secondary school students, we had some thoughts regarding how to relate this artifact and Ubimus. Our analysis is based on the two forces of attraction proposed by Keller and Barreiro (2018) in Ubimus' researches. We believe that the artifact has potential to fit into both forces of attraction.

Regarding the first force of attraction, everyday creative practices, the artifact presented here can support formal music education practices, such as the music learning process and the building of musical discourse in the music classroom. The artifact's use in the school environment, as evidenced by the fun and pleasant characteristics demonstrated on the focus group and observed in the classroom and in the recorded video, shows that the artifact also has the potential to support musical practices that do not have pedagogical intention or design, inside and outside the school. In the second case, there are opportunities for future works that can also investigate the mingling of different user profiles when it comes to musical experience.

Regarding the second force of attraction and more specifically its third aspect, the influence of ubiquitous computing techniques in the development of Ubimus, we have other considerations about the new relationships between musical instruments and those who play them.

Since the invention of the first musical instruments, such as flutes made out of bones or drums carved in trunks or in the ground, human beings had to adjust to the acoustic characteristics of their environment and, especially, to the physical characteristics of the musical instrument chosen to intervene in this environment. It was necessary to create an instrument to translate the sound intentions of a subject into sound vibrations that altered the environment, be it through the vibration of strings, membranes or air columns. With time, the techniques of instrument production developed, but some of the mechanic-acoustic characteristics of instruments were impossible to change. For example, it would be impossible to build an ukulele that sounded like a bass and produced the same notes. Sound making and sound performance technique are confined to the physical characteristics of acoustic instruments (such as shape and constituting materials), because these characteristics physically interact with the environment in order to produce sounds. There are several postural consequences of the relationship between the body of the player and the musical instrument they are playing, and human beings, for a long time, were dependent on the physical characteristic of instruments in order to learn how to use their bodies to activate them. Thus, the body of the person conforms to the body of the instrument. In order for this to occur, the player, whether they are a professional or a layperson, has to understand how the instrument of their choice produces sound, and also learn the necessary technique to produce sounds with that instrument.

The advent of synthesizers and computers allowed different timbres from different instruments to be played in a single instrument (keyboard), but the player still had to adjust to the shape of this single instrument.

Therefore, in our understanding, Ubimus presents new possibilities of interaction between the body of the user and the instrument and, also, regarding the ways in which users collaborate. With Ubimus techniques, we can adjust the digital instrument to the characteristics of the user's body, inverting the ergonomic relationships between player and instrument. We can also adjust the artifact to the environment where the activities will occur. As we showed with the Smart Musical Mats, it is possible to decrease technical barriers in sound making and promoting collective interaction in sound making, creating more alternatives for building musical discourse. The access of laypeople to music becomes wider, and music becomes something that can be democratically shared not only through its execution, but also through the possibility of building one's own musical discourse.

The article written by Brown & Ferguson (2024) includes great expectations of popularizing the process of building sound instruments with electronic components, custom-made PCBs and microcontrollers and processors using the possibilities of the DIY approaches in Ubimus. However, these approaches are based on sound instruments which, although they can be hacked, reprogrammed and built by a public with a knowledge of electronics and of the use of certain digital manufacturing technologies, make certain ergonomic demands of their users in the sense of not exploring body movements more intensely during the performance. In a way, they constitute other options of using the fingers and the hands to connect with sounds and, even though there are some innovations in the manner in which sounds are synthetized using microcontrollers and other bespoke hardware, they do not seem to facilitate new body instrument relationships during sound production. The article written by Merendino (2024) deals with this issue of the body by presenting two instruments that can modify sound aspects as a response to other physical gestures. Both the Dispositivo Cinetico Midi and the Sonic Cubes use user gestures that go beyond finger movement, so characteristic of acoustic instruments such as the piano and the flute, in order to influence sound result.

Another issue presented in these articles published in the Journal of Ubiquitous Music (2024) is the gradual extinction of the line that used to separate the instrument builder from the performer that used the instrument.

Ubiquitous computing technologies, digital manufacturing technologies and maker culture applied in DYI initiatives in Ubimus have contributed to the emergence of the Designer-Performer, overcoming the dichotomy present in the case of acoustic instruments with the Luthier and the Player. When I designed and built the Smart Musical Mats, I also had in mind creating a sound instrument that I could explore with my elementary school students, supporting their musical education. Therefore, I became a different kind of centaur, one with three parts: a third musician, a third teacher and a third designer. After all, what led me to this were musical, pedagogical and technological issues. And these issues feed each other in constant dialectics in the process of investigating and building the Smart Musical Mats.

8 CONCLUSION

Our research problem is directly linked to our first conjecture about decreasing technical barriers in sound making to support musical literacy. In order to do that, we developed an artifact at the same time that we developed our research, in which we used the Design Science Research approach. In our evaluation of the results, we noted that the decrease of technical barriers in sound making occurred, and that students were able to build their musical discourse without much effort. Therefore, we consider that our problem was solved. From now on, we believe that it is necessary to investigate how the artifact performs in other contexts besides year 7 classes of municipal schools in Rio de Janeiro, so that we can have more trust in the tendencies presented here.

Regarding the acceptance of the SMM by students, we observed that the artifact has acceptance potential and that its use in different contexts by users with different profiles can serve as a basis for future investigation. We feel it is important to note that the proposed artifact does not mean to be a substitute for traditional musical instruments, but an alternative to increase the options available for students to build musical discourse in the school environment.

The artifact also encourages collaboration between students and the use of their bodies and body movements, as well as the bodies and body movements of others in building musical discourse. We shaped the artifact to adjust it to the users' bodies, but this artifact can also be adjusted according to the characteristics of the environment where the activity will occur, such as a classroom. This fact also creates developments in the ways in which users collaborate while experimenting with ubiquitous computing musical systems such as SMM (Santos, 2015).

The possible physical configurations in the arrangement of the artifact in the classroom allow users to move freely and observe the movements of other participants, which increments interactions and the formulation of hypotheses about the building of musical discourse. Since sound making with the artifact does not demand a great deal of cognitive effort, users can pay attention to their environment and exchange information with other users, be it they verbal or gestural.

More importantly, the decrease of technical barriers in sound making allows the understanding and application of musical concepts, and allows the communicational intentions and expressions of each subject to be channelled and concretized through the building of their musical discourse.

LIMITATIONS AND FUTURE WORKS

One limitation present in this version of the artefact is that the user, while making sounds, cannot alter the volume of the available sound samples. This way, musical issues regarding dynamics cannot yet be explored in this version of SMM and may be investigated in future researches.

A possible line of future investigation is the use of visual recognition to detect gestures to give inputs in sound parameters such as the volume of each sound produced with the artefact or even to allow the triggering of sound samples.

Another possible line of investigation is passing the sound processing from a software in a computer inside the classroom to a website and being able to connect to it several smart objects besides the mats, with the web becoming part of the Internet of Things as investigated in Santos (2023).

REFERENCES

Brown, Andrew; Ferguson, John. DIY musical instruments: From Handmade Electronic Circuits to Microcontrollers and Digital Fabrication. Journal of Ubiquitous Music, 2024.

Dalcroze, E.J. (1921) Rhythm music and Education.G.P.Putnam’s Sons New York.

Dresch, A., Lacerda, D. P., & Júnior, J. A. V. A. (2015). Design science research: método de pesquisa para avanço da ciência e tecnologia. Bookman Editora.

Frazee, J; Kreuter, K (1997) Discovering Orff: A Curriculum for Music Teachers (Schott) Paperback – January 1997.

Hevner, A; Chatterjee, S. (2010) Design Research in Information Systems. Ed. Springer New York 2010

Keller, Damián & Schiavoni, Flávio & Lazzarini, Victor. (2019). Ubiquitous music: Perspectives and challenges. Journal of New Music Research. 48. 1-7. 10.1080/09298215.2019.1659828.

Keller, D., & de Lima, M. H. (2018). O enfoque cognitivo-ecológico no design ubimus– Da Metáfora da Listra aos Sons do CAp. Revista Vórtex, 6(2).

Keller, D., & Barreiro, D. L. (2018). Ubiquitous music section: Forces of attraction and challenges in ubimus research (editorial). Revista Vórtex , 6 (2), 1-14.

Keller, D., Aliel, L., & Silva, C. R. (2018). The handy metaphor: Bimanual, touchless interaction for the internet of musical things. In Anais do workshop em música ubíqua / proceedings of the ubiquitous music workshop (ubimus 2018). São João del Rei, MG: Ubiquitous Music Group

Keller, D., & Lazzarini, V. (2017). Ecologically grounded creative practices in ubiquitous music. Organised Sound, 22(1), 61-72.

Keller, D., Capasso, A. (2006). New concepts and techniques in eco-composition. Organised Sound 11 (1), 55-62. (Doi: 10.1017/S1355771806000082.)

Lima, M. H., Keller, D., Flores, L. V. (2018). Eco-composition and everyday creative musical practices: theory and practice experience in ubiquitous music research at UFRGS Application School with highschool students in and out of the classroom. In: MUSICHILDREN'17. Music for and by Children:online proceedings of the international conference musichildren'17. [ISBN: 978-972-789-539-7]. Aveiro, Portugal [:s.n.]. p.139-151.

López, Tomás Sánchez, et al. "Taxonomy, technology and applications of smart objects." Information Systems Frontiers 13.2 (2011): 281-300.

Makey Makey (2020) http://www.makeymakey.com/ Latest access: 02/20/2020

Merendino, N. (2024) DIY Musical Instruments and Communities: From Handmade Electronic Circuits to Microcontrollers and Digital Fabrication, v. 1 n. 1, p. 35-41.

Miletto, E. M.; Pimenta, M. S.; Bouchet, F.; Sansonnet, J.-P.; Keller, D. (2011) Principles for music creation by novices in networked music environments. Journal of New Music Research, v. 40, n. 3, p.205-216.

Nicolaci-da-Costa, A. M. (2007). O campo da pesquisa qualitativa e o Método de Explicitação do Discurso Subjacente (MEDS). Psicologia: reflexão e crítica, 20(1), 65-73.

PARÂMETROS CURRICULARES NACIONAIS. (PCNs) (1998) 2. Arte : Ensino de quinta a oitava séries. Secretaria de Educação Fundamental. – Brasilia : MEC / SEF.

Pimenta, M & Flores, L & Kuhn, C & Farias, F & Keller, D & Lazzarini, V. (2015). Música Ubíqua: Suporte para atividades musicais em dispositivos móveis. ScientiaTec. 2. 61-74. 10.35819/scientiatec.v2i2.1469.

Pimentel, M; Filippo, D; SANTOS, T. M. Design Science Research: pesquisa científica atrelada ao design de artefatos. RE@ D-Revista de Educação a Distância e Elearning, v. 3, n. 1, p. 37-61, 2020.

Pimentel, M., Filippo, D., Calvão, L.D., & Silva, A. R. (2017). “Design Science Research: pesquisa científica para o desenvolvimento de artefatos inovadores”. In: Pesquisa e Inovação: visões e interseções. editado por R. M. Araujo & L. O. V. Chueri. Rio de Janeiro, RJ: Publit.

Richards, R., (ed.). (2007) Everyday Creativity and New Views of Human Nature: Psychological, Social, And Spiritual Perspectives. Washington, DC: American Psychological Association,. (ISBN: 978-0-9792125-7-4.)

Santos, T.M. (2023) CiberTRIBOrgue: Experiências Ciborgues Apoiando a Educação Musical. Tese de Doutorado. Programa de Pós-Graduação em Informática (PPGI), Federal University of the State of Rio de Janeiro (UNIRIO), Rio de Janeiro. 2023.

Santos, T. (2015) Computação Ubíqua para apoiar a Educação Musical: Explorações com Tapetes Musicais Inteligentes. 212 pgs. Dissertação de mestrado - Programa de Pós Graduação em Informática (PPGI), Federal University of the State of Rio de Janeiro (UNIRIO), Rio de Janeiro. 2015.

Stolfi, A., Milo, A., & Barthet, M. (2018). Cannibal soundscape [ubiquitous music work]. In Anais do workshop em música ubíqua / proceedings of the ubiquitous music workshop (ubimus 2018). São João del Rei, MG: Ubiquitous Music Group.

Serafine, M. L. (1988). Music as Cognition: The Development of Thought in Sound. Columbia University Press.

Sloboda, J.A. (1997) In Wilson, Frank & Roerman, Franz L. Music and child development. St. Louis, MO: MMB Music Inc.

Simon, H. A. (1969) The Sciences of the Artificial (1st ed.), Cambridge, MA: MIT Press.

Soares, M. (2004). Letramento e alfabetização: as muitas facetas. Revista brasileira de educação, (25), 5-17.

Soundplant (2020) <http://soundplant.org/> Latest access: 02/20/2020.

Swanwick, K. (1979) A basis for music education. London: Nfer,. England.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS quarterly, 425-478.

Vygotsky, L. (1989) Pensamento e linguagem. São Paulo: Martins Fontes, 1989.

Weiser, M. (1991) The computer for the twenty-first century. Scientific American, 65(3):94-104

Yin, R. K. (2015). Estudo de Caso-: Planejamento e métodos. Bookman Editora.

MAUDLIN: INTERFACE FOR RANDOM WALKS AND BEYOND - EUROCK MODULE FOR UBIQUITOUS MUSIC

JOHN HARDING[1]

J.HARDING@ULSTER.AC.UK

ABSTRACT

This paper offers an expanded review of work presented at the ubimus 2025 conference, detailing the MAUDLIN Eurorack module inspired by the “drunk” object in Cycling 74’s Max/MSP. MAUDLIN introduces a novel interface for generating controlled randomness in ubiquitous music (ubimus) practices by leveraging an unconnected Analog to Digital Converter (ADC) pin as a natural entropy source to augment pseudo-random sequences constrained within user-defined boundaries. Designed for modular synthesisers, the module integrates control voltage (CV) inputs, potentiometers, and an display for real-time interaction and feedback, aligning with ubimus principles of accessibility and creative flexibility. We detail the module’s de sign principles and the use of environmental (EMF) noise to seed and perturb randomisation, we also discuss the potential to enhance musical creativity through constrained random walk outputs. As an open-source hardware and software development, MAUDLIN aims to democratise access,

---

[1] Ulster University - Derry/Londonderry

fostering collaborative innovation in the ubimus and Eurorack communities while serving as a potential platform for expanded applications in embedded technology contexts.

1 INTRODUCTION

Ubiquitous music (ubimus) emphasises accessible, context-aware, and collaborative musicmaking, leveraging everyday technologies to foster creativity [Keller et al., 2014]. Randomness has long been a creative force in music, from the aleatoric compositions of John Cage to modern generative systems that introduce constrained unpredictability [Barton and Fritz, 2024]. Constrained randomness, where random processes operate within defined boundaries, balances serendipity with intentionality, enabling novel outcomes to be accessible to diverse practitioners. In ubimus, such approaches lower barriers to musical expression, aligning with its ecologically grounded ethos of creativity emerging from human-technology-environment interactions.

The MAUDLIN Eurorack module aligns with these principles by providing a compact and interactive tool for generating controlled randomness in modular synthesis. Inspired by the ’drunk’ object in Max / MSP [Cycling ’74, 2025], which implements a Markov chain random walk [Konstantopoulos, 2009], MAUDLIN extends this concept to the hardware domain, offering a tactile interface for live performance and composition. A key innovation is centred around the use of an unconnected ADC pin as a source of environmental (EMF) noise, capturing inherent electrical fluctuations to augment the production of constrained random sequences. This approach offers a supplementary approach to digital pseudo-random number generators (PRNG) [Knuth, 1998] or the more traditional analog transistor-based white noise circuits [Moog Music Inc., 1978], providing a system that enhances musical expressivity [Barton and Fritz, 2024].

This paper presents MAUDLIN’s design, implementation, and relevance to ubiq uitous music practice, emphasising its role in democratising creative tools through open source development and its capacity to foster intuitive, unpredictable but constrained musical interactions. By bridging analog and digital domains, MAUDLIN aims to serve as a model for future embedded system creativity, with potential applications beyond music in generative art and networked systems [Turchet et al., 2018].

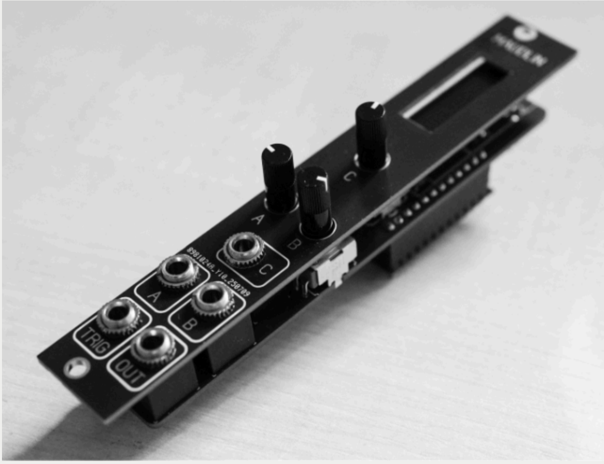


FIGURE 1: MAUDLIN HARDWARE V 1.0

2 DESIGN AND IMPLEMENTATION

The MAUDLIN unit in standard Eurorack format, depicted in Figures 1 and 2 and authored by John Harding in 2025, is specified as follows:

2.1 ARCHITECTURE

The core functionality of this development is inspired by the random walk functionality of the ’drunk’ object [Cycling ’74, 2025], where the values are produced in response to a trigger and drift within a user-defined range, and where the current value is probabilistically related to those that preceded it, reflecting the dynamic processes of analog computers [Lazzarini and Timoney, 2020]. An Arduino Pro Mini microcontroller [Arduino, 2025] is selected for its low cost, small footprint, and ease of use. An external clock/trigger input allows synchronisation, and three control voltage (CV) inputs labeled; A, B and C allow for modulation via external CV sources while associated potentiometers; A, B and C allow for manual adjustment. Finally, a Digital to Analog (DAC) converter produces the CV output [Microchip Technology Inc., 2009].

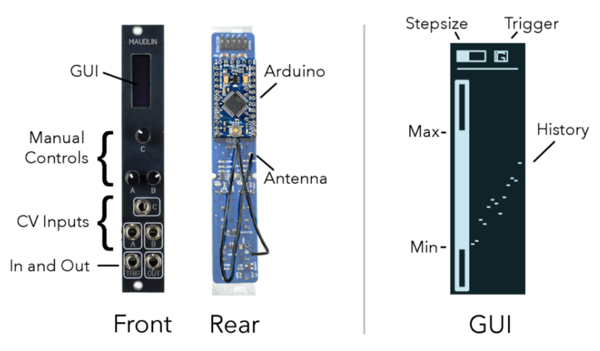


FIGURE 2: MAUDLIN HARDWARE V 1.0 AND GUI

The unconnected or floating ADC pin with an optional attached antenna, shown in Figure: 2, serves as an entropy source, capturing electronic/analog environmental (EMF) noise to seed the random walk in terms of directionality and step size, which is con strained within user-defined boundaries and uses reflection to maintain continuity similar to the nonlinear dynamics of analog synthesis systems. [Lazzarini and Timoney, 2020]. The interface prioritises accessibility, allowing users to intuitively implement controlled randomness without extensive technical knowledge, a core ubimus tenet.

While initial focus is placed on the Markov chain random walk application under current discussion, the hardware is intentionally designed to facilitate a series of other functions in the future, ensuring that MAUDLIN is adaptable for various musical uses and settings. The features allow the hardware to be used for various optional modes in future, adding to the versatility across diverse musical contexts.

2.2 HARDWARE SPECIFICATION

To conform to the Eurorack ethos of minimal horizontal width (HP), MAUDLIN is de signed with surface-mount electronic components, occupying 8 HP. It adheres to Eurorack power constraints (±150mA per ±12V rail)[1]. Components include the Arduino Pro Mini, MCP4725 DAC, SSD1306 OLED, MCP6002 rail-to-rail op-amps for CV input and output conditioning, and 3.5mm jacks for CV and gate inputs and one CV output. The front panel features a cutout for OLED visibility with labeled controls. The interface prioritises accessibility, allowing users to shape randomness intuitively. [Keller et al., 2014]. Full technical details are available through the online repository: [Harding, 2025a].

3 FIRMWARE

The firmware processes inputs to generate a random walk seeded by EMF noise, scaled and directed within user-defined boundaries, and interfaced with an MCP4725 Digital to- Analog Converter (DAC) [Microchip Technology Inc., 2009] for CV output. The initial design focuses on a range of 0V to 5V DC output CV, with future development planned for ±5V output via an open-ended conditioning circuit. The SSD1306 OLED [Solomon Systech Limited, 2012] provides real-time feedback of control parameters and output values. The system integrates CV inputs (step size, range, offset) and corresponding potentiometers, and the module adheres to Eurorack standards [Doepfer Musikelektronik, 2023].

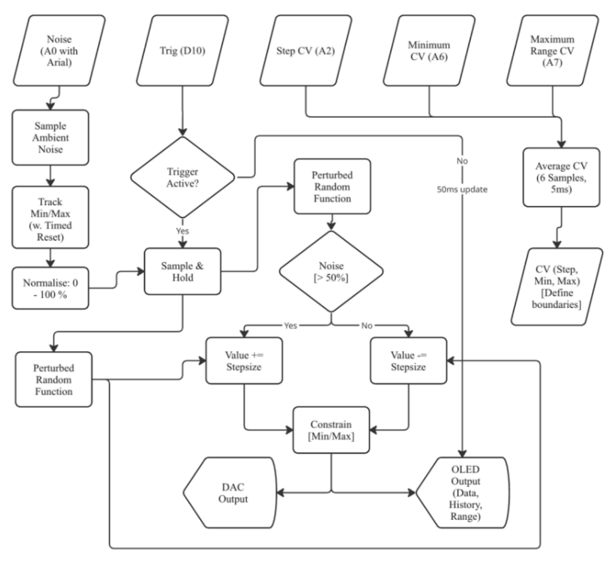


FIGURE 3: MAUDLIN: DRUNK INSPIRED, MARKOV-CHAIN. SIMPLIFIED LOGICAL FLOW DIAGRAM

3.1 FIRMWARE: LOGICAL FLOW AND STRUCTURE

As depicted schematically in Figure: 3, environmental (EMF) noise is captured via an unconnected or floating ADC pin (A0), with an optional 5-10 cm antenna attached to the rear

---

1. Doepfer Eurorack Standard [Doepfer Musikelektronik, 2023]

of the PCB, serving as an entropy source. A double read protocol (discarding the first A0 read after grounding A1, with a 10μs delay) minimises crosstalk between adjacent channels. On a trigger input (via D10), the noise is sampled and held, with normalisation to a 0-100 percent range using dynamically tracked min/max boundaries, reset every 1000 samples (~5 s) following a 20-sample (~100 ms) stabilisation period. For low noise variation (range <50), the normalised value is amplified (multiplier = 100 / range). The normalised noise percentage offsets a PRNG base-direction value, constrained to 0-100%, to determine the step direction. Step magnitude is a similarly perturbed random value from 0 to stepsize and if EMF readings are below 20, the system offers a fallback mechanism to pure PRNG to ensure reliable performance across lower measured EMF environments. This augmentation process blends PRNG reliability with noise entropy for robust randomness across varied EMF environments, with the walk value constrained to [minimum, maximum] before DAC output. A 128x32 SSD1306 OLED displays a step size bar, min/max slider, output history, and gate indicator, updated every 50 ms. The trigger-to-output response time ranges from 2.7-3.4 ms, supporting rates up to 294 Hz, with nonblocking C++ routines ensuring low-latency operation. Alternate firmware modes may support increased rates depending on the coding complexity.

3.2 FIRMWARE BIAS AND AUTO-ADJUST

The latest Drunk V.5 firmware iteration[2] enhances the Markov chain random walk with bias correction and auto-adjustment to optimise noise distribution for a less biased

[FormattingFormatting]

ran domisation of step size and directionality. The module retains its core structure: CV inputs (STEP_CV, OFFSET_CV, RANGE_CV) control step size (0-512), minimum (0-4095), and maximum (0-4095), respectively, with noise from NOISE_PIN pertubing direction and step size via a normalised 0-100% value. A gate trigger (TRIG_PIN) updates the DAC output, constrained within user-defined boundaries, and the OLED displays a step bar, slider, history, and gate indicator (Figure: 2). Key features include:

- **Bias Adjustment:** An autoBias variable (initially 30) is applied to the normalised noise percentage post-mapping, adjusting the 0-100% range to correct for skew (e.g., shifting more values above 50% for positive steps). This is constrained between 0 and 50 to prevent overcompensation.

- **Auto-Adjust Mechanism:** The system calculates a running mean of raw (pre-bias) percentNoise over 100 held samples. If the mean deviates beyond a 5% hysteresis band (below 45% or above 55% from the target 50%), autoBias increments or decrements by 1, ensuring the distribution centers dynamically.

- **Robustness:** Full initialisation of averaging buffers to 512, a watchdog timer resetting after 8 seconds, and improved min/max resets (falling back to 100/900 if invalid) enhance reliability. [FormattingFormatting]

2.https://bitbucket.org/Nonzerojohn/maudlin/src/main/Maudlin%20Code/

- Performance: Non-blocking timing (5ms sampling, 50ms display) and ADC crosstalk mitigation (GND_PIN resets, double-read) maintain low latency, suitable for real time Eurorack use.

This solution, detailed further in the following sections, preserves the original intent while adapting to environmental noise variations.

3.3 NOISE AUGMENTATION FOR VARIABLE ENVIRONMENTS

The development of the MAUDLIN Drunk inspired firmware highlighted challenges in utilising environmental electromagnetic (EMF) noise as an entropy source to drive the Markov chain random walk. A series of tests were conducted which revealed a bias towards negative steps due to the limited dynamic range (~0-250 ADC units) and frequent zero values, resulting from its lower end skew during normalisation to a range of 0- 100% via dynamically tracked min/max values. This constrained variability, tested to vary across environments, prompted a hybrid solution. To mitigate inherent biases—such as skew toward lower values observed in initial testing—the system employs dynamic min/max tracking and periodic resets, ensuring the noise distribution adapts to varying conditions. Hybrid techniques further enhance reliability: pseudo-random number generation (PRNG) [Knuth, 1998] is augmented with true noise, which seeds both step-size and directionality, blending determinism with organic fluctuation. The aforementioned auto-adjustment mechanism monitors the mean of normalised noise values, incrementally tuning bias (initially set at 30) to center the distribution around the midpoint, promoting balanced entropy without over-reliance on either source. This approach addresses practical challenges like zero-dominance and, if noise measurements fall below 20, the system falls back to utilising pure PRNG for both step size and direction. This ensures consistent performance across varied EMF environments: noise adds subtle perturbations in low-variation environments and dominates in high, blending sources for robustness across contexts.

3.4 EXAMINATION OF NOISE CONDITIONING: THE UPDATENOISE(); FUNCTION

To ensure a reliable entropy source, the updateNoise() function processes raw noise data from the unconnected ADC pin (A0), adapting to environmental electromagnetic field (EMF) fluctuations. The relevant function is presented below:

```
void updateNoise() {
  noiseValue = rawNoise;

  // Update min/max tracking
  if (noiseSampleCount >= STABILIZATION_SAMPLES) {
    if (rawNoise < minNoise) minNoise = rawNoise;
 if (rawNoise > maxNoise) maxNoise = rawNoise;
 } else {
 noiseSampleCount++;
 }
 resetCounter++;
 if (resetCounter >= RESET_INTERVAL) {
 minNoise = (rawNoise > 0) ? rawNoise : 100;
 maxNoise = (rawNoise < ADC_RESOLUTION) ? rawNoise : 900;
 resetCounter = 0;
 }
}
```

This function initialises noiseValue with the latest raw ADC reading and implements a stabilisation period (STABILISATION_SAMPLES = 20) before updating the min/max boundaries. Once stabilised, it tracks the minimum and maximum noise values dynamically, ensuring the normalisation range reflects current conditions. The reset mechanism, triggered every 1000 samples (~5s), re-initialises these boundaries with fall backs (100/900) if the current noise is invalid, preventing stale distributions. This adaptive approach mitigates biases (e.g., zero-skew) observed in initial tests. Future optimisations could involve exponential moving averages for smoother tracking, though this increases computational overhead on the 16MHz Pro Mini.

3.5 EXAMINATION OF EMF PERTURBATION MECHANISM

The EMF perturbation mechanism in MAUDLIN's Drunk firmware represents an integration of environmental entropy into a controlled random walk. As shown graphically in Figure 3: the perturbation is applied to both the direction and step size, using the bias adjusted and normalised noise percentage (0-100%) derived from the ADC pin (A0). For direction, the base PRNG value (random(0, 101)) is offset by (percentNoise - 50), restricted to 0-100, allowing EMF noise to subtly influence the probability of posi tive or negative steps. For step size, the base random value (random(0, stepsize + 1)) is perturbed by a noise-derived offset, calculated as map(percentNoise, 0, 100, -stepsize/10, stepsize/10), and constrained to 0-stepsize. This scaling ensures the perturbation remains proportional to the step size, typically ±51 for a maximum stepsize of 512, adding organic variation without exceeding bounds.

Example: Direction Perturbation

The base PRNG direction (random(0, 101)) is offset by percentNoise - 50, constrained to 0-100, allowing EMF noise to bias the probability of positive or negative steps.

```
int baseDirection = random(0, 101); // PRNG base for direction int
augmentedDirection = baseDirection + (percentNoise - 50); // Noise
augments (offsets) PRNG
augmentedDirection = constrain(augmentedDirection, 0, 100);
```

Example: Step Size Perturbation

The step size is calculated using random(0, stepsize + 1) and per turbed by a noise-derived offset, scaled via map(percentNoise, 0, 100, -stepsize/10, stepsize/10). This keeps perturbations proportional (e.g., ±51 for stepsize=512), constrained to 0-stepsize.

```
int randomStep = random(0, stepsize + 1); // Random step 0 to stepsize
int noiseOffset = map(percentNoise, 0, 100, -stepsize/10, stepsize/10);
randomStep = constrain(randomStep + noiseOffset, 0, stepsize);
```

3.6 EXAMINATION OF FALLBACK AND AMPLIFICATION

[Indent]The implementation includes a fallback mechanism for low noise ranges (< 20), which mirrors the direction process. If currentRange < LOW_NOISE_THRESHOLD, both direction and step perturbations revert to pure PRNG values, ensuring functionality in environments with minimal EMF interference (e.g., 0-19 range). Amplification stretches noise variation for ranges 20-50, tested to handle transitions from low to high noise (0- 260+), validated through oscilloscope measurements showing consistent step distribution. This hybrid approach balances computational efficiency with environmental randomness, critical for adaptability.

[Indent]For EMF low noise ranges (<20), the mechanism falls back to pure PRNG for both direction and step size, ensuring reliability in low-EMF conditions (e.g., 0-15 range). For ranges 20-50, amplification stretches noise variation, validated by oscilloscope tests showing consistent step distribution across high-noise transitions (0-260+).

Example: Fallback Mechanism

```
 int currentRange = maxNoise - minNoise;
 if (currentRange < LOW_NOISE_THRESHOLD) {
 rawPercent = random(0, 101); // Pure PRNG for percent }else if
(currentRange < MIN_NOISE_RANGE && currentRange >=
LOW_NOISE_THRESHOLD) {
  int amplification = 100 / currentRange;
  rawPercent *= amplification;
  rawPercent = constrain(rawPercent, 0, 100);
}
```

3.7 FIRMWARE PERFORMANCE AND OPTIMISATION

MAUDLIN's performance on the Arduino Pro Mini (16MHz) has been tested in situ to meet the demands of real-time Eurorack and music performance, with a focus on trigger rates up to ~250 Hz[3], aligning with the predicted speeds. The 5ms sampling interval, tailored to the ADC's conversion time [4], supports rapid random walk updates, processed in a non-blocking loop to minimise jitter. The MCP4725 DAC's I2C write (~100µs), deferred with a non-blocking flag, contributes to an estimated trigger response of approximately 130µs, as inferred from previous oscilloscope tests across noise ranges (0-15 to 0-260 ADC units). This response enables a maximum trigger rate of ~250 Hz (4 ms period), suitable for both dynamic live modulation and rhythmic applications. In real-world use, MAUDLIN has been informally tested in casual studio sessions and small performances, where the 50ms OLED update interval provided clear visual feedback on step size, range, trigger status, and history without any noticeable lag. Estimated performance under continuous triggering at 250 Hz suggests stable operation, with perturbation calculations (e.g., noise offset for steps) adding a negligible ~2-3 µs (estimated) to the overhead, well within a 4 ms cycle. Field observations in varied EMF environments (e.g., varied indoor settings) indicate that hybrid randomisation adapts effectively.

4 ALTERNATE FIRMWARE MODES

To demonstrate MAUDLIN's versatility as a module with scope beyond the primary application discussed in this paper, several alternate firmware modes are currently under development, utilising the same hardware for various creative tasks. For instance, a quantiser mode maps input CV to selectable scales (e.g. chromatic, major, minor, pentatonic,

---

3. See Video Reference 7: [Harding, 2025b]
4. ADC conversion time for the ATmega328P is approximately 104 µs at a 16MHz clock with a 128 prescaler, as specified in the datasheet [Atmel Corporation, 2020]

22-tone equal temperament) with CV control over root-note, note trigger and scale form, displaying input sliders and output history. A digital LFO, LFOyeah, produces 12 user-created waveforms at CV-controlled rates and shapes, ideal for modulation purposes. Finally, a Random Source mode samples noise on trigger, normalising to full DAC range for direct output, with graphical sliders and scrolling history. These modes, while in early stage development at the time of writing, hope to underscore the module's adaptability, enabling users to re-purpose it for CV generation, intonation, modulation, or random value generation by simply changing the firmware, all applications aim to maintain the low-latency and open-source extensibility aspects and will be delivered via the code repository.

5 OPEN-SOURCE APPROACH

The firmware, written in C++ for Arduino via Platformio IDE[5] along with hardware design files are presented open source [Harding, 2025a] to encourage community use and contributions, aligning with the ubimus emphasis on collaborative creativity. Users may modify or extend functionality, such as adding alternate firmware. The universal nature of CV and Gate inputs, along with 0 - 5 VDC CV output, enables broader applications in modular synthesis. The hardware development presented through this work has potential for greater scope and future development beyond the Eurorack format and towards standalone applications. At the time of writing, multiple alternative

[FormattingFormatting]

---

[5] https://platformio.org/

firmware applications are in development and will be located within the aforementioned code repository once completed. [Harding, 2025a]

6 MUSICAL APPLICATIONS AND UBIMUS RELEVANCE

MAUDLIN has the potential to enhance ubimus practices by providing a hardware-based tool for generative music that is relatively low-cost, accessible, and adaptable. The current random-walk algorithm, seeded and augmented by ambient electrical EMF noise, aids in production of evolving patterns which are constrained within user defined boundaries. These parameters can be set manually via potentiometers and/or modulated in real time via Control Voltage inputs for Step Size, Minimum, and Maximum boundaries. The organic but controlled unpredictability contrasts with deterministic digital systems, enriching the sonic palette for musical applications. The tactile interface and causally linked visual feedback enable musicians to explore randomness intuitively. [Keller et al., 2014].

[Indent]We contend that constrained randomness of this nature significantly reduces barriers to entry for musical creation, allowing novices to produce musically compelling results without requiring extensive technical or musical expertise [Kramann, 2020]. This approach aligns with the principles of ubimus, which emphasise accessible and ecologically grounded creative frameworks [Keller et al., 2014]. This development aims to facilitate the creation of harmonic structures that evolve over time, offering musically engaging outputs that are accessible to users with minimal training.

[Indent]Kramann's work provides some parallels, illustrating how structured, game-like systems with constrained randomness can guide lay participants in musical improvisation [Kramann, 2020]. With a single trigger input, and when coupled with a following quantisation module, a constrained random walk can be used to generate musically coherent CV pitch sequences by producing complete melodic passages for a VCO, with OLED feedback for accessibility; this mirrors Kramann's structured improvisation by empowering novices to create engaging music through simple game-like interactions.

[Indent]The principles of exploration take inspiration from the nonlinear interactions implicitly seen in the Buchla-Serge synthesis paradigm, supporting the idea of generative systems that produce unpredictable yet musically meaningful outcomes, as seen in modules like the Serge Dual Universal Slope Generator [Barton and Fritz, 2024] or Smoothed Stepped Generator [Serge Modular Music Systems,] which by their very nature encourage emergent sonic outcomes.

[Indent]The open source design invites community-driven enhancements and alternative future applications, such as generative quantisation to 22-tone equal temperament for example [Erlich, 1998], while integration with networked platforms like the Internet of Musical Things [Turchet et al., 2018], may expand its accessibility in the future. Using low-cost embedded hardware, this has the capacity to serve as a model for future embedded systems, with potential applications in generative art, interactive installations, or educational tools [Keller and Lazzarini, 2017] and [Lazzarini and Timoney, 2020]. This democratisation of advanced music technologies supports the vision of collaborative,

[FormattingFormatting]

everyday musical creativity.

6.1 COMPARISON WITH TRADITIONAL EURORACK PATCHING

It is challenging to implement a completely equivalent patch using a series of standard Eu rorack modules; however, as a suggested approximation to this goal and with a comparative range of main features, it would require as a minimum:

- Make Noise Wogglebug (10 HP)[6] – Provides noise, random CV (for steps), and clocked randomness.

- Intellijel Quadrax (14HP)[7] – Handles Sample and Hold on trigger, gate input, and can generate variable steps via envelope modulation.

- Doepfer A-172 Min/Max (4 HP)[8] – Analog min/max selector to clamp the output. Output directly as CV.

[Indent]This approach would require significantly more horizontal width (HP) than the current development. As an approximation; 28 HP. This solution would not provide the user with any form of intuitive visual feedback and is a costly and bulky solution. We would argue that the visual feedback component is vital for the user in this case. The 8 HP design with integrated OLED feedback is more compact and accessible, aligning with the expectations of the Eurorack standard, where module horizontal width (HP) is a key [FormattingFormatting]

---

6. https://www.makenoisemusic.com/modules/wogglebug/
7. https://intellijel.com/shop/eurorack/quadrax/
8. https://www.doepfer.de/a100_man/A172_man.pdf

factor in determining commercial acceptance whilst maintaining the mobility of the system overall.

7 VIDEO DEMONSTRATIONS

A video playlist showcasing the core and extended functionality of MAUDLIN has been provided to expand on the technical descriptions provided in this extended paper and to provide further detail the practical applications of this development for musical uses. These demonstrations highlight the core functionality, accessibility, and musical usage in the context of Eurorack synthesis, and in the future this playlist will serve as a placeholder for demonstrations of alternate firmware modes as they are publicly released. These video references playlist can be found at the following location: [Harding, 2025b].

8 CONCLUSION AND FUTURE WORK

We believe that MAUDLIN represents a significant contribution to both the ubiquitous music community and the Eurorack synthesis ecosystem by merging analog randomness with digital control and GUI in a compact, low-cost

9 ACKNOWLEDGEMENTS

This work is supported by Ulster University, DLD Campus, N.Ireland.

REFERENCES

Arduino (2025). Arduino pro mini. https://www.arduino.cc/en/Main/ArduinoBoardProMini. Accessed: 2025-05-09.

Atmel Corporation (2020). 8-bit AVR Mi crocontroller with 32K Bytes In-System Programmable Flash: AT mega328/P Datasheet. Microchip Technology Inc. Available: https://ww1.microchip.com/downloads/en/DeviceDoc/Atmel-7810-Automotive-Microcontrollers-ATmega328P_ Datasheet.pdf.

Barton, T. and Fritz, I. (2024). Randomness and uncertainty in sound design. In Filimowicz, M., editor, The Routledge Handbook of Sound Design. Routledge.

Cycling ’74 (2025). Drunk reference. https://docs.cycling74.com/max8/refpages/drunk. Max 8 Documentation, Accessed: 2025-05-09.

Doepfer Musikelektronik (2023). Technical details a 100: Construction details. https://doepfer.de/a100_man/a100m_e.htm. Accessed: 2025-18-08.

Erlich, P. (1998). Tuning, tonality, and twenty-two tone temperament. Xenharmonikôn, 17. Revised 4/2/02, An Informal Journal of Experimental Music.

Harding, J. (2025a). Maudlin repository. https://bitbucket.org/Nonzerojohn/maudlin/src/main/. Bitbucket Repository, Accessed: 2025-09-05.

Harding, J. (2025b). Maudlin video repos itory. https://www.youtube.com/playlist?list=PLvcKlQ9XRG3WAhMlwhTqmWXe6nu7HDsR4. Video Repository, Accessed: 2025-11-08.

Keller, D. et al. (2014). Ubiquitous music: Perspectives and chal lenges. Journal of New Music Research, 43(1):1–6.

Knuth, D. E. (1998). Seminumerical algorithms. In The Art of Computer Programming, pages 10–26. Addison-Wesley, third edition.

Konstantopoulos, T. (2009). Markov chains and random walks. https://www2.math.uu.se/~takis/L/McRw/mcrw.pdf. Lecture Notes, Accessed: 2025-11-08.

Kramann, G. (2020). Of renouncing to do something grandiose. In Stolfi, A., Costalonga, L., Messina, M., Keller, D., and Aliel, L., editors, Proceedings of the 10th Workshop on Ubiquitous Music (UbiMus 2020), pages 21–35, Porto Seguro, BA. UFSB.

Lazzarini, V. and Timoney, J. (2020). The analogue com puter as a musical instrument. In Ubiquitous Music. Taylor & Francis.

Microchip Technology Inc. (2009). Mcp4725 12-bit digital-to-analog converter with eeprom memory in sot-23-6. Accessed: 2025-05-09.

Moog Music Inc. (1978). Minimoog service manual. https://www.synthfool.com/docs/Moog/minimoog/Minimoog_Service_Manual.pdf. Figure 9-6, page 100, Accessed: 2025-11-08.

Serge Modular Music Systems. Smooth & stepped function generator. https://sdiy.info/wiki/CGS_Serge_smooth_and_stepped_generator. Accessed: 2025-08-11

Solomon Systech Limited (2012). Ssd1306 128 x 64 dot matrix oled/pled segment/common driver with controller. https://cdn-shop.adafruit.com/datasheets/SSD1306.pdf. Rev 1.1, Accessed: 2025-05-09.

Turchet, L., Fischione, C., Essl, G., Keller, D., and Barthet, M. (2018). Internet of musical things: Vision and challenges. IEEE Access, 6:61994–62017.

STANDALONE INTERACTIVE AND GENERATIVE MUSIC WITH THE CSOUND-FPGA FRAMEWORK

AMAN JAGWANI [*1] · **VICTOR LAZZARINI** [1]

AMAN.JAGWANI.2023@MUMAIL.IE **VICTOR.LAZZARINI@MU.IE**

ABSTRACT

This paper presents and discusses approaches to interactive and generative music composition and deployment on Field Programmable Gate Array (FPGA)-based System on Chips (SoCs) with the Csound-FPGA framework. With the development of Csound 7, Csound is now able to target bare-metal embedded systems, opening up the possibility of running Csound on SoCs that contain both programmable hardware and processing systems. This provides a unique platform for deploying standalone generative musical systems, leveraging the flexibility and computational power of the FPGA alongside the portability of an established domain-specific language like Csound on the CPU. With careful planning and division of tasks between the Programmable Logic (PL) and the Processing System (PS), comprehensive systems and compositions can be depoloyed, in contrast with other embedded systems like microcontrollers that may be more constrained. This paper presents an overview of the framework along with an examination of how it can contribute towards ubimus interactive

---

1. Department of Music, Maynooth University - Maynooth, Ireland.

music practices. Following that, a complete generative and interactive musical system is presented as an example study, highlighting possible methodologies for deploying this framework. Lastly future possibilities, limitations and conclusions are discussed.

1 INTRODUCTION

With Csound 7, the ability to target bare-metal embedded platforms has been introduced [Lazzarini and Jagwani, 2025]. One such platform capable of running Csound is the field programmable gate array (FPGA)-based System-on-Chip (SoC). These chips typically consist of FPGA fabric, or programmable logic (PL), integrated with an ARM CPU or processing system (PS) that can run bare-metal software. The Csound-FPGA framework involves running bare-metal Csound on the PS, along with custom audio processing hardware modules on the PL. The framework manages all low-level detailes and handles communication between the PS and PL.

---

1. Department of Music, Maynooth University - Maynooth, Ireland.

This paper presents an overview of the Csound-FPGA framework and explores the opportunities it offers for generative and interactive music design within ubimus contexts. Generative music systems are typically self-contained, adaptive, and algorithmically driven, capable of evolving autonomously over time [Gradim and Pestana, 2021]. Randomness, probability, stochastic systems and flexible event scheduling are often key elements in shaping their behavior. When applied to a diverse palette of timbral and textural sound sources, these strategies can result in a complete generative music system.

[Indent]While such systems can be built using sample-based or fixed media sources, digital signal processing (DSP) offers additional possibilities by enabling the dynamic shaping and synthesis of sound. In this way, DSP can serve as a powerful tool in generative music design. Given the availability of both generative and DSP features, domain-specific, high-level music programming environments such as Csound [Lazzarini et al., 2016] are particularly well suited to support these approaches.

[Indent]Generative music also implies the fact that while musi-cians, artists or composers define musical parameters or instructions such as note choice, sequencing rates and timbral possibilities, they are not part of the actual musical or performance process, with the generative system independentaly realising the musical outcome at the time of performance. This independence of performance can suggest that standalone or embedded deployments may be particularly well-suited to support the self-evolving, self-

[FormattingFormatting]

---

*Aman Jagwani would like to acknowledge the support of the Hume Scholarship Programme, Maynooth University.

contained nature of such systems.

However, due to the complex logic, sequencing structures, and signal processing often required to sustain musical and timbral richness, generative systems are commonly deployed on desktop platforms, where computational resources are abundant. This can be seen in works such as [Mangwani, 2024]. Yet, this reliance on desktop environments can limit portability. Embedded platforms, while more compact and deployable, often lack the computational resources or DSP capabilities required for such tasks. For instance, ubimus toolkits based on microcontrollers such as the Interactive Audio Toolkit (IAT) [Jagwani and Lazzarini, 2025] offer flexibility and musical interfaces for sensor-based interactions and mechanical sequencing, but may be limited in audio processing capacity.

FPGAs can address this gap by offering a reconfigurable embedded platform that supports both DSP and generative audio. As field-programmable devices, FPGAs align well with the concept of aesthetic pliability [Keller et al., 2023], allowing hardware configurations to be shaped and reshaped for specific artistic needs. Moreover, this adaptability connects with the idea of the Internet of Musical Stuff [Messina et al., 2024], where”musical stuff” is defined as morphable and context-driven. With support for both wired and wireless networking, as well as the ability to run web-servers on FPGA SoCs as demonstrated in [Murugan et al., 2016], FPGAs are well positioned to function as part of this environment.

Being embedded also allows FPGAs to create standalone systems that integrate directly with a wide range of sensors and peripherals, enabling interactivity to complement generative music. This opens new possibilities for audience engagement. For example, an FPGA could autonomously generate an evolving soundscape in an installation, while real-time data from LiDAR or ultra-sonic sensors modulate aspects of the audio, such as spectral brightness, spatialisation, or dynamics, based on audience movement. This kind of interaction allows for meaningful participation by non-expert audiences, aligning with the ubimus concept of inclusivity [Keller et al., 2019b].

Such systems offer a promising platform for inclusive interaction design, accommodating both passive and active engagement. They can support the kind of embodied participation seen in the Chaal installation discussed in [Jagwani and Lazzarini, 2025], and can extend the ideas explored in the Handy Metaphor [Keller et al., 2019a], where touchless, sensor-based musical interactions were facilitated using technologies like ultrasonic and vision-based systems. FPGAs, with their ability to handle sensor data and perform tasks such as video-based computer vision [B.K. et al., 2017], can extend these approaches.

One common barrier associated with FPGAs is their relative inaccessibility, particularly due to the complexity of hardware design and low-level programming [Jagwani, 2023]. However, the Csound-FPGA framework abstracts these low-level details and introduces Csound as a high-level interface for working with the FPGA platform. This opens up the technology to a broader community of musicians and composers who may not have experience with hardware design but are comfortable with musical programming environments, once again highlighting the idea of inclusivity but in the realm of the composer or artist in this case instead of the audience.

[Indent]With this improved accessibility, the framework can ali-gn with the second wave of ubimus, particularly with the ideas of do-it-yourself (DIY) practices [Keller et al., 2024]. Timoney et al. [Timoney et al., 2020] outlined the evolution of DIY ubimus practices, examining microcontroller-based DIY ubimus platforms like Arduino. The Csound-FPGA framework can build on this foundation by offering a more powerful, yet still accessible, platform for DIY exploration, particularly due to its integration with a familiar music programming language.

[Indent]Additionally, [Brown and Ferguson, 2024] discussed di-gital fabrication as a continuation of DIY ubimus practices, following the transition from non-programmable to programmable electronics. Similarly, the move from general-purpose CPUs to custom, programmable hardware via FPGAs represents a further step in this progression. This transition is only possible, however, if the tools involved provide a high enough level of access, something the Csound-FPGA framework facilitates. Moreover, while digital fabrication enables the creation of custom circuit boards, this framework mirrors this idea with the designing of custom hardware circuits directly in the FPGA's programmable logic (PL), effectively bringing a new layer of fabrication to DIY musical practices.

[Indent]As highlighted in [Keller et al., 2025], "possibly one of the most complex challenges faced by ubimus frameworks is the demand for supporting both legacy practices and prospective exploratory initiatives." The Csound-FPGA framework is particularly well suited to address this challenge. Csound, with its history as a widely used FLOSS environment, brings strong connections with legacy musical practices while maintaining an active and evolving user base. At the same time, FPGAs, with their reconfigurable, morphable hardware, embody the spirit of exploratory design, offering new possibilities for signal processing, interaction, and generative systems. Thus, this framework can align well with the ubimus concepts of replicability and aesthic pliability while being conducive to innovation and exploration in a cutting-edge technological platform.

[Indent]The subsequent sections provide background on the use of FPGAs in audio and music applications. This is followed by an overview of the Csound-FPGA framework, including its mechanisms, methodologies, and how it supports common generative music practices within Csound. The paper concludes with a complete example application demonstrating generative and interactive music design using the framework, along with a discussion of potential future developments.

2 FPGAS IN AUDIO AND MUSIC

FPGAs are integrated circuits with programmable logical units that can be configured into custom hardware through low-level hardware design techniques. Along with their inherent flexibility and reconfigurability, in an audio processing context, FPGAs provide the benefits of ultra-low latency, high-throughput, high sampling rates and comprehensive connectivity through a large number of general-purpose inputs and outputs (GPIOs) [Jagwani, 2023]. These features make them well-suited to musical tasks.

[Indent]Commonly, FPGAs are packaged as a [FormattingFormatting] system-on-chip with an associated CPU (PS). A C or C++ application generally runs on the PS while custom hardware runs on the FPGA fabric or PL. This system architecture provides the opportunity for efficient division of tasks across both parts of the SoC through hardware-software co-processing. The hardware used for experimentation and discussion in this paper was the Xilinx Zynq 7000 SoC, running on a Digilent Zybo Z7020 development board [Digilent, 2023].

[Indent]Generally, FPGAs are complex to program. Low-level hardware design languages such as Verilog and VHDL are most commonly used to program them. These may be out of reach for the general music and audio community, making FPAGA usage and adoption for DIY ubimus practices and instrument design limited.

[Indent]However, recent explorations of higher-level methodo-logies and frameworks for FPGA programming have provided wider access to the platform. An early exploration of the use of high-level synthesis (HLS) for FPGA-based audio synthesis was seen in [Fitzgerald, 2019], highlighting how HLS can enable hardware designs to be generated from C or C++ programs. [Popoff et al., 2023] presented the Syfala toolchain which enables complete FPGA applications to be designed with the FAUST [GRAME, Centre National de Cr´eation Musicale, 2025] language. [Keller et al., 2025] presented a ubimus plugging framework in the form of ModFPGA, highlighting a system for modular audio synthesis and processing on FPGAs.

[Indent]All of these studies present comparatively user-friendly forms of FPGA based audio programming but their

[FormattingFormatting]

focus was generally on the technical aspects of running audio processing programs. Musical and compositional approaches are not extensively discussed. Additionally, these studies show FPGA audio programs being generated from higher-level sources but not actually running a complete higher-level domain-specific language like Csound itself on an FPGA SoC platforms.FPGA SoC platforms.

[Indent]These gaps can be addressed with the Csound-FPGA framework, particularly in the context of standalone generative and interactive music design.

3 THE CSOUND-FPGA FRAMEWORK

The FPGA SoC architecture necessitates a careful division of tasks between the PS and PL, with control-rate operations typically assigned to the PS and audio-rate processing handled by the PL, as demonstrated in [Keller et al., 2025, Popoff et al., 2023].

[Indent]With the development of Csound 7 and the emergence of its co-processing capabilities, enabled by bare-metal csound [Lazzarini and Jagwani, 2025], it is now possible to run Csound on the PS while delegating audio processing operations to custom intellectual property (IP) cores or modules in the PL. [Keller et al., 2025] presents the methodologies and strategies for deisgning these IP cores with HLS in detail. These can include custom hardware modules or hardware ports of Csound opcodes such as reverbsc [Lazzarini and Jagwani, 2025]. The overall system architecture is summarised below:

[FormattingFormatting]

- **PS-based processing:** The PS runs bare-metal Csound, handling scheduling, sequencing, modulation, and control tasks, as well as potentially performing some synthesis and audio processing.

- **PL-based processing:** High-throughput, low-latency audio-rate processing is performed on the PL. These operations take place on a sample-by-sample basis and can be run at very high sampling rates, enabling the exploration of techniques such as higher-order FM synthesis [Lazzarini and Timoney, 2024] or virtual analog synthesis.

- **PS-PL communication:** The PS and PL communicate in different ways depending on the type of data they are sharing. For control data, the AXILITE protocol [AMD, 2024] is used and for the streaming audio data, a Direct Memory Access (DMA) IP/hardware module is used. More details for the communication methodologies are presented in [Lazzarini and Jagwani, 2025].

- **MIDI input:** MIDI can be received via USB or through a UART peripheral using one of the Zybo board's assignable IO pins.

- **Sensor input:** Analog sensors are connected to the system via the XADC Wizard IP core [AMD, 2012] in the PL. The data from this IP core can be accessed from the PS with the XADC driver and is transferred to Csound through its sofware bus [The Csound Developers, 2025] or control channels.

- **Additional connectivity:** The Zybo board also facilitates additional connectivity; a networked system can be integrated with the help of the ethernet port or a visual system can be integrated with the help of the HDMI port on the board, for example.

- **System Extension:** The Zynq chip contains a large amount of flexible, assiginable general-purpose inputand output (GPIO) pins, which can be assigned a range of protocols such as UART, I2S or SPI, enabling flexible communication with other devices, breakout boards or even other microcontrollers. For example an ESP32 board running the Interactive Audio Toolkit (IAT) described in [Jagwani and Lazzarini, 2025] can be connected to a Zybo board running the Csound-FPGA framework. This can allow the ESP32 to handle sensor-based interactions, WI-FI connectivity and mechanical output sequencing while the Zybo handles generative audio and DSP tasks. Combined with the DIY digital fabrication process presented in [Brown and Ferguson, 2024], a custom PCB for a standalone extended hybrid interactive sound installation system can be created.

[Indent]Thus, the Csound-FPGA Framework, takes care of task delegation between the PS and PL, and handles communication protocols and interactivity. With these lower level details abstracted, the user can simply interact with the framwork like any other Csound project. For instance, the Csound main engine input and output buffers, spin and spout are automatically connected to a DMA IP that manages audio data communication with the on-board

[FormattingFormatting]

audio CODEC. Users can simply call the ins, outs opcodes for audio IO. For using a hardware processing module on the PL, users can simply use the plug-in opcodes, fpgaOut and fpgaIn for sending and receiving audio data as explained in [Lazzarini and Jagwani, 2025]. Internal connections and DMA transfers are once again, taken care of by the framework. The system architechture is illustrated in figure 1.

4 GENERATIVE MUSIC WITH CSOUND ON FPGAS

Csound presents vast opportunities for generative music design through flexible sequencing capabilities with opcodes such as sequ, random number and probabilistic generators such as rspline, jspline, random, randi, randh as well as flexibile scheduling with opcodes like schedkwhen. The code excerpt below shows how some of these tools can be utilised to create a simple generative melodic system. Instrument 1 is the main clocking system that triggers Instrument 2, which is the melodic instrument based on FM synthesis, at continuously evolving rates. When triggered, Instrument 2 chooses a random note from gkseq, which is an array of midi note numbers.

```
gkseq[ ] fillarray 71, 67, 79, 74, 71, 74, 79, 76, 83, 76

instr 1 // generative clocking system
krandcps rspline 0.1, 4, 0.1, 0.5
gkrandtrig metro krandcps
kranddiv trandom gkrandtrig, 1, 6
schedkwhen gkrandtrig, 0, 1, 2, 0, 0.1 + rnd(0.3)
endin
```

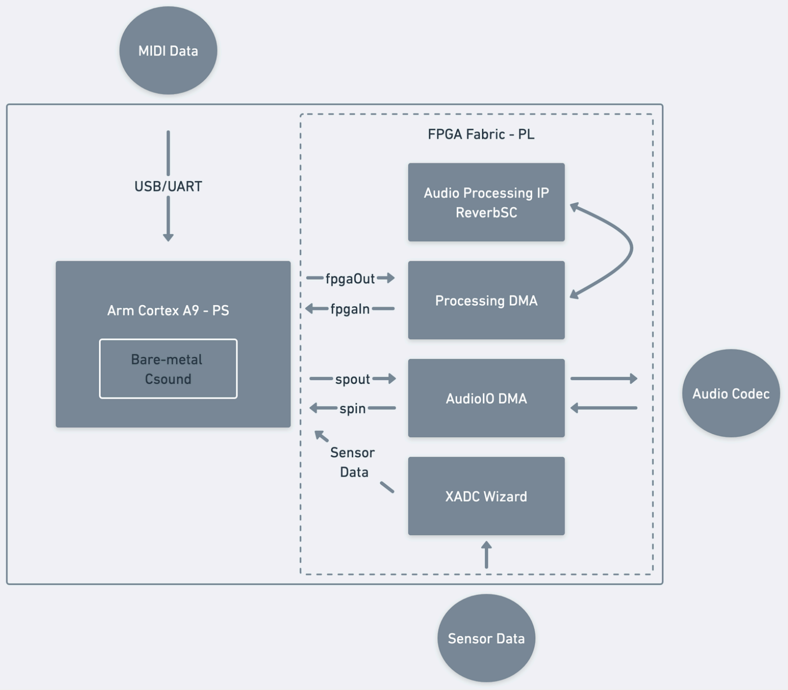


FIGURE 1: CSOUND-FPGA FRAMEWORK ARCHITECTURE

```
instr 2 // FM mallet-style instrument
index random 0, 10
index = int(index)
imodindx = rnd(1)
knote = gkseq[index]
iamp = (0.7 + rnd(0.3))
idur = p3
knote = cpsmidinn(knote-24)
aindx expseg 30, idur/2, 0.00001
aenv expsegr 0.001, 0.01, 0.1, idur, 0.01, idur/4, 0.001
amod1 oscili (aindx*imodindx)*(knote*4), knote*3
acar1 oscili 1, knote+amod1
aout = (acar1)*aenv*iamp
outs aout, aout
endin

schedule(1, 0, 1000)
```

An installation that utilises Csound for its generative so-und design can be seen in [Mangwani, 2024]. In this installation, all sounds are generated in real-time with Csound running on a desktop computer. Additionally, heart rate and breath sensors are used to create modulations in sound and lights in coherence with a participant's physiological responses. This installation illustrates the power of combining generative techniques with interactive systems. The generative piece is able to evolve autonomously while the breath and heart-rate provide a level of grounding and connection with the audience, allowing them to participate in a seemingly passive but meaningful way. This also enables the extension of audience participation into passive, physiological domains, increasing inclusivity and allowing non-expert audiences to contribute to the musical process without requiring active technical engagement.

However, a key limitation of this installation is its relian-ce on a desktop platform. In general, target platforms for generative music tend to be confined to Desktop and possibly web platforms. This may be due to the lack of resources in commonly available embedded platforms to run complex generative and audio processing tasks efficiently. Furthermore, embedded programming often requires low-level C or C++ development, which may be limiting for composers. While platforms like Daisy [Electro-Smith, 2023] and Arduino [Ard,] do simplify some aspects, they are limited in providing high-level, expressive generative capabilities. Another limitation of this installation is the reliance on proprietary, commercial devices such as the Fitbit smart watches [Google LLC, 2025] for heart-rate and breath sensing. These bring additional costs and limitations of vendorsupplied APIs for accessing data.

[Indent]The Csound-FPGA framework can provide a unique and suitable solution for this. The availability of a large amount of logical resources in the PL along with the ability to divide tasks between the PS and the PL, means that complex generative programs can easily be run in an embedded, portable context. The embedded FPGA SoC format also allows on-board interfacing with sensors, removing the need for any proprietary devices, reducing costs and providing flexibility in real-time data access. For example, the Analog Devices MAX30102 [Analog Devices, 2016] can replace the Fit-bit in the above installation. At the same time, the familiarity and all of the generative abilities of Csound are preserved, allowing composition and sound design at a higher, more comfortable level for artists. Additionally, the portability of Csound programs, ideas and sketches across platforms also means that the FPGA-based SoC can become a platform to explore and deploy previous Csound developments such as the Organic Generative Structures presented in [Heintz, 2024].

5 EXAMPLE

This section presents an example of a generative and interactive music piece with the Csound-FPGA Framework.

[Indent]In this example, sequencing, synthesis, and audio genera-tion are all handled by Csound running on the PS, while audio processing, specifically with the hardware port of the reverbsc opcode, is handled by the PL. This approach reduces the computational load on the PS by delegating a complex feedback delay network to the FPGA, which

[FormattingFormatting]

processes it on a sample-by-sample basis with ultra-low latency. As a result of this offloading, the system can operate with a lower ksmps value in Csound, minimising latency. Retaining synthesis on the PS also enables the use of Csound’s unique opcodes, such as gbuzz, which is featured in this example. Two sensors are used for adding modulations, a Li-DAR sensor along with a light dependent resistor (LDR). Both of these add motion-based interaction to the system with the LiDAR sensor reacting to long-range motion and the LDR reacting to close-range motion.

[Indent]The code below contains the csound program used in this example, highlighting the use of the software bus for sensor data and FPGA plugin opcodes for PL processing:

```
gkseq1[] fillarray 71, 67, 74, 79, 74, 71, 74, 79
gkseq2[] fillarray 76, 71, 79, 83, 79, 76, 79, 83
gisine ftgen 0,0,4096,10, 1

instr 1 // sequencer
ktempo chnget "lidar" // LiDAR sensor modulating tempo
gkharms chnget "ldr" // ldr sensor modulating harmonic content
gkbuzz metro ktempo
gkpad metro 0.125*ktempo*0.5
gkdur = 1/(ktempo)
kadditions[] fillarray -12, 0, 12, 0, -12, 0, 12, -12
kindx trandom gkbuzz, 0, 8
kindx = int(kindx)
knote = gkseq1[kindx]
knote2 = gkseq2[kindx]
```

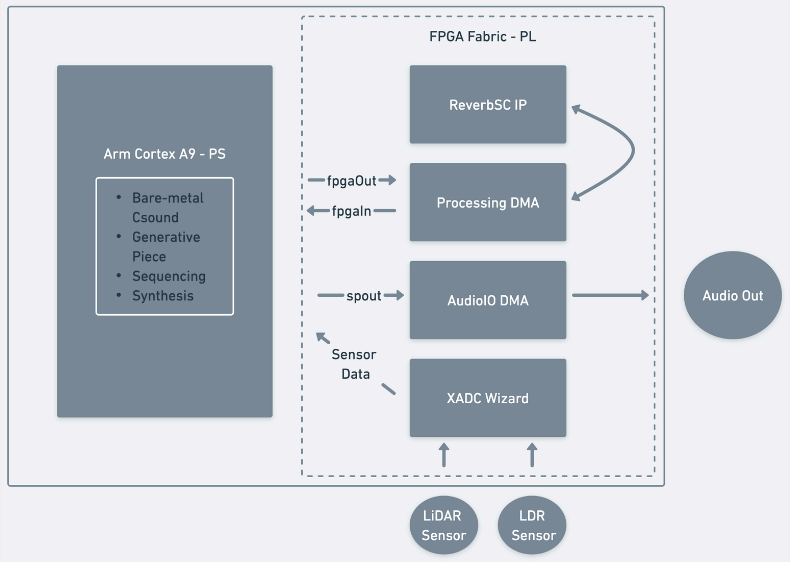


FIGURE 2: EXAMPLE GENERATIVE AND INTERACTIVE SYSTEM DESIGN WITH THE CSOUND-FPGA FRAMEWORK

```
kadd = kadditions[kindx]
schedkwhen gkbuzz, 0, 8, 11, 0, gkdur*0.5, knote, 0.3, 0
schedkwhen gkpad, 0, 8, 11, 0, 0.25*16, knote2+kadd-12, 0.3, 1
endin

instr 2
aenv linsegr 0, p3*0.5, 0.4, p3*0.25, 0.6, p3*0.25, 0.4, p3*00.25, 0
kdetune1 rspline 0.01, 3, 0.1, 3
kdetune2 rspline 0.01, 3, 0.1, 3
kmul rspline 0.1, 0.99, 0.01, 5
kamp rspline 0.1, 0.4, 0.01, 0.1
a1 gbuzz aenv*kamp*(0.4+(0.4)), cpsmidinn(p4-12)+kdetune1, 50, 1, gkharms *
kmul, gisine
a2 gbuzz aenv*kamp*(0.4+(0.4)), cpsmidinn(p4-12)+kdetune2, 50, 1, gkharms *
kmul, gisine
kpan rspline 0, 1, 0.1, 2
al, ar pan2 a1 + a2, kpan
```

```
// FPGA opcodes used to send audio to PL for processing
fpgaOut al, ar
afpgaL, afpgaR fpgain
// outs connected to audio output from Zybo board
outs (afpgaL)*p5, (afpgaR)*p5
endin

schedule(1, 0, 10000)
```

This example can be heard at the link below:
https://youtu.be/sEWzhKcb3wc

[Indent]The system design and delegation of tasks is highlighted in figure 2.

[Indent]While this example is effective sonically and musically, the hardware-software configuration can also be modified to have Csound function primarily as a control engine, handling sequencing and parametric control on the PS with all audio synthesis and signal processing tasks implemented in the PL, using modular IP blocks from the ModFPGA ubimus plugging framework [Keller et al., 2025]. This demonstrates how the Csound-FPGA framework can integrate with another ubimus platform seamlessly for system design. With all of the audio processing in the PL, this alternate configuration could further leverage the FPGA’s capabilities of low-latency as well as its computational resources.

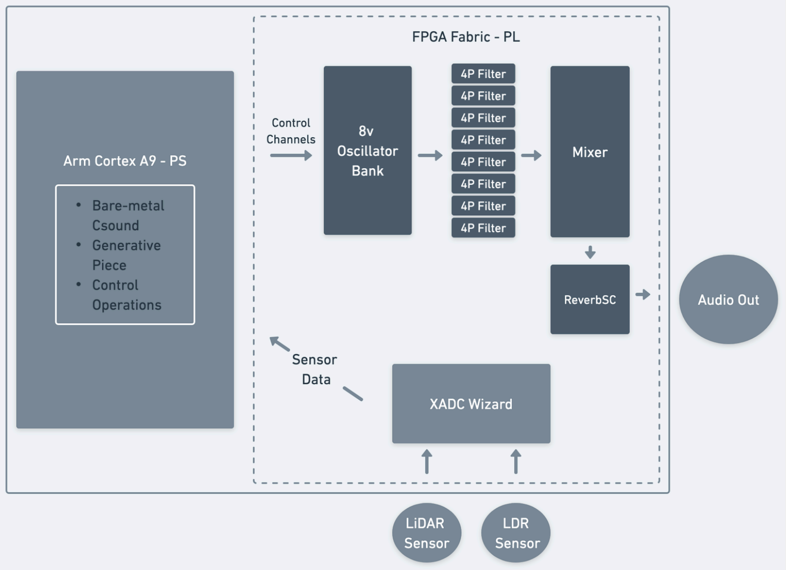


FIGURE 3: ALTERNATIVE GENERATIVE AND INTERACTIVE SYSTEM DESIGN WITH THE CSOUND-FPGA FRAMEWORK INTEGRATING WITH MODFPGA IP CORES

For example, a synthesizer with eight polyphonic voices with two oscillators and a fourpole filter each, along with the reverbsc module can comfortably run on the PL on a sample-by-sample basis and Csound can generatively control this synthesizer from the PS.

[Indent]This alternate system design can be seen in figure 3.

6 FUTURE DEVELOPMENTS AND CONCLUSIONS

The Csound-FPGA framework can present a compelling platform for interactive ubimus practices, combining the strengths of both Csound and FPGAs. While FPGAs are inherently flexible, the integration of a complete and mature library like Csound makes artistic deployment even more pliable through custom Csound code. This is further supported by the ability to leverage different hardware–software co-processing configurations within the same system, as demonstrated in the example discussed.

[Indent]Importantly, Csound itself is not a new tool, but a well-established and familiar environment now extended to a new platform. This continuity can help lower the learning curve and reduce barriers to entry, enabling the portability of existing Csound projects to this relatively new embedded ubimus platform with greater ease.

[Indent]Furthermore, the combination of ease of use, computa-tional power, and the inherently interactive ability of embedded systems offers a unique space for generative DSPbased composition that also supports real-time interaction. This opens up new possibilities for engaging non-expert audiences in the music-making process, allowing them to participate meaningfully through intuitive, self-adapting and responsive systems that do not require technical engagement.

[Indent]Future work would include the development of additio-nal Csound opcode ports as hardware IP modules. In particular, FPGA parallelism can be leveraged for computationally intensive techniques such as spectral processing or time-varying convolution. Another key development would be the automation of the build process, enabling users to generate complete FPGA applications directly from .csd files. While the current framework abstracts many low-level details, it still involves a considerable setup process, particularly with respect to the Xilinx tool chain. With these improvements, the framework can evolve into a more comprehensive system for DSP, interactivity, and generative audio design, offering a more a more accessible entry point for musicians and artists.

[Indent]The Csound-FPGA framework, along with the example discussed in this paper, will be available in the Csound github repository:

https://github.com/csound/csound

REFERENCES

Arduino. https://www.arduino.cc/. Accessed: 2024-02-16.

AMD (2012). LogiCORE IP XADC Wizard v2.2 User Guide (UG772). AMD Adaptive Computing. Version 2.2.

AMD (2024). MIPI D-PHY LogiCORE IP Product Guide (PG202): AXI4-Lite Interface. AMD Adaptive Computing. Version 4.3.

Analog Devices (2016). MAX30102: High-Sensitivity Pulse Oximeter and Heart-Rate Sensor. Analog Devices. Accessed: 2025-05-10.

B.K., A., Venkatraman, V., Kumar, A. R., and S., S. D. (2017). Accelerating real-time computer vision applications using hw/sw co-design. In 2017 International Conference on Computer, Communications and Electronics (Comptelix), pages 458–463.

Brown, A. R. and Ferguson, J. (2024). Diy musical instruments: From handmade electronic circuits to microcontrollers and digital fabrication. Journal of Ubiquitous Music, 1(1):8–22.

Digilent (2023). Zybo z7 reference manual. https://digilent.com/reference/programmable-logic/zybo-z7/start. Accessed:2023-08-24.

Electro-Smith (2023). Daisy. https://www.electro-smith.com/daisy. Accessed: 2023-06-08.

Fitzgerald, L. (2019). Sound synthesis using programmable system-on
[FormattingFormatting]

chip devices. Master’s thesis, National University of Ireland, Maynooth, Co. Kildare, Ireland.

Google LLC (2025). Google store watches. Accessed: 2025-05-10.

Gradim, R. and Pestana, P. D. (2021). Overview of generative processes in the work of brian eno. In Proceedings of the 11th Workshop on Ubiquitous Music (UbiMus 2021), pages 45–56, Matosinhos, Portugal. HAL ID: ffhal-03398725f.

GRAME, Centre National de Cr´eation Musicale (2025). Faust: Functional audio stream. Accessed: 2025-05-14.

Heintz, J. (2024). Creating organic generative structures in csound. In Stojak, S. and Hofmann, A., editors, Proceedings of the 7th International Csound Conference, Hannover, Germany. ICSC. Accessed: 2025-05-14.

Jagwani, A. and Lazzarini, V. (2025). Interactive audio toolkit: Creating sonic experiences and installations with low-cost, low-power microcontrollers. In Estadieu, G. V., Messina, M., Gomez Meja, C. M., Keller, D., Kramann, G., and Koszolko, M., editors, Proceedings of the Ubiquitous Music Symposium 2024 (UbiMus 2024), pages 1–132, Macau. Ubiquitous Music Symposium.

Jagwani, A. P. (2023). Developing a modular sound synthesis platform for fpgas with high level synthesis techniques. Master’s thesis, Maynooth University, Department of Music.

Keller, D., Gomes, C., and Aliel, L. (2019a). The handy metaphor: Bimanual, touchless interaction for the internet of musical things. Journal of New Music Research, 48:1–12.

Keller, D., Jagwani, A., and Lazzarini, V. (2025). The ubimus plugging framework: Deploying fpga-based prototypes for ubiquitous music hardware design. Computers, 14(4):155.

Keller, D., Lazzarini, V., Turchet, L., and Brooks, A. L. (2024). Ubimus contributions to digital creative practices (editorial). Digital Creativity, 35(1):1–12.

Keller, D., Messina, M., Bridges, B., and Yaseen, A. (2023). Editorial: Ubiquitous music as a movement. In Proceedings of the Ubiquitous Music Symposium 2023 (Ubimus 2023), Derry, Ireland. Ubiquitous Music Group. HAL Id: hal-04372736.

Keller, D., Schiavoni, F., and Lazzarini, V. (2019b). Ubiquitous music: Perspectives and challenges. Journal of New Music Research, 48:1–7.

Lazzarini, V., Ffitch, J., Yi, S., Heintz, J., Brandtsegg, O., and McCurdy, I. (2016). Csound: a sound and music computing system. Springer.

Lazzarini, V. and Jagwani, A. (2025). CSOUND 7: Bare Metal and Co-Processing Hardware. In Proceedings of the 19th Linux Audio Conference (LAC-25), Villeurbanne, France.

Lazzarini, V. and Timoney, J. (2024). Theory and practice of higher-order frequency modulation synthesis. Journal of New Music Research, 0(0):1–16.

Mangwani, H. (2024). Effection: A biofeedback-based interactive art installation. https://www.hansikart.com/biofeedback-art. Accessed: May 10, 2025.

Messina, M., de Souza Stolfi, A., Aliel, L., Simurra, I., and Keller, D. (2024). The internet of musical stuff: Towards an aesthetically pliable musical internet. International Journal of Software Innovation (IJSI), 12(1):1–19.

Murugan, V. D., Saravanan, C., and Imtiaz, R. (2016). Fpga based standalone embedded web server for remote control of display devices. International Journal of Microelectronics and Embedded Systems, 4(2):224–228.

Popoff, M., Michon, R., Risset, T., Cochard, P., Letz, S., Orlarey, Y., and de Dinechin, F. (2023). Audio DSP to FPGA Compilation: The Syfala Toolchain Approach. Technical Report RR-9507, Univ Lyon, INSA Lyon, Inria, CITI, Grame, Emeraude.

The Csound Developers (2025). Csound Manual: Software Bus. Csound Project. Accessed: 2025-05-10.

Timoney, J., Lazzarini, V., and Keller, D. (2020). Diy electronics for ubiquitous music ecosystems. In Lazzarini, V., Keller, D., Otero, N., and Turchet, L., editors, Ubiquitous Music Ecologies, chapter 3. Routledge.