Interstices - Zwischenräume

An interactive sound installation: Design, implementation, evaluation

Toningenieursprojekt

Raphael Kapeller

Betreuung: Mag.art. Mag.rer.nat. David Pirrò

Graz, November 29, 2017

kunst uni graz

institut für elektronische musik und akustik

iem

Abstract

This project report examines the process of designing, implementing, staging and evaluating the interactive sound installation *Interstices - Zwischenräume*. The goal was to create an installation that alters the acoustic environment, providing various and changing acoustic responses, either imitating real world acoustics or evoking completely new auditory perspectives. As the installation reacts to sound, visitors were encouraged to walk through the installation space and to use the provided objects to produce sound, in order to change the state of the installation and eventually the whole acoustic environment. During the designing phase, different algorithms and tools were developed, scenarios were designed and explored, and an interaction pattern based on physical modelling was realized and fine-tuned. We will document the exploratory process, presenting the encountered problems and the applied solutions. Further, this work documents the results of an evaluation of the visitors' experiences using grounded theory.

Zusammenfassung

In dieser Arbeit soll gezeigt werden, wie die interaktive Klanginstallation *Interstices - Zwischenräume* geplant, umgesetzt, inszeniert und evaluiert wurde. Ziel war es, eine Installation zu schaffen, die durch unterschiedliche, veränderliche akustische Signale Einfluss auf den Raum nimmt. Dabei wurden zum einen reale akustische Umgebungen imitiert und zum anderen völlig neue akustische Perspektiven eröffnet. Da die Installation auf Schall reagiert, wurden die BesucherInnen dazu motiviert, sich frei im Ausstellungsraum zu bewegen und die bereitgestellten Objekte zu verwenden, um Geräusche zu erzeugen. Dadurch veränderte sich der Zustand der Installation und somit die akustische Umgebung. In der Planungsphase wurden verschiedene Algorithmen und Anwendungen entwickelt, Szenarien konzipiert und untersucht und ein Interaktionsmuster auf Basis der Physikalischen Modellierung umgesetzt und genau parametrisiert. In dieser Arbeit sollen der Untersuchungsprozess und die auftretenden Probleme und Lösungsmöglichkeiten dokumentiert werden. Zudem werden unter Anwendung der Grounded Theory die Erfahrungsberichte von BesucherInnen analysiert.

Contents

1 Introduction

A sound installation is an intermedia and time based art form. [Wik] It differs from other installation types by presenting sound as the main part of the art form. Sound vanishes, as long as it is not stored on a media or as long as it can not be reproduced. The constant of time gives the possibility, to use sound in order to interact with the environment. By producing sound purposefully or accidentally, information about the surrounding can be collected by listening to the given feedback. This process works invisibly as a background task. Only in new, unexpected, not definable or unfamiliar situations, where visual, auditive or expected perceptions do not fit together, it is shifted to come to the fore.

With *Interstices - Zwischenräume* we wanted to develop a sound installation, that invites visitors to be part of the installation space. Visitors should be able to form the installation responses ranging from subtle details to flamboyant outcomes, depending on their behaviour and their attentional focus. In this work we will outline the development of the installation covering the conceptual approach, technical design, final staging and perception evaluation.

This work is classified into three major parts. The first part gives a general description about the staging of the installation. It describes general aspects like the concept of the installation, the scenery or the used objects (Chapters 1.1; 1.2). The second part examines and describes the technical aspects of developing an interactive sound installation. The focus lies on giving information about the applied theoretic approaches, implementation and problem solving. Also, artistic and aesthetic decision influence the programming and they will be highlighted to allow a further understanding of the creative process. (Chapters 2; 3) The third part is about visitors and how they perceived the sound installation. In order to gather some knowledge about an installation visit and about the experiences made during the visit, a theory using *grounded theory research* is developed (Chapter 5). Finally, the work is concluded in Chapter 6.

1.1 Concept of the installation

Interstices - Zwischenräume is an interactive sound installation, that alters the acoustic environment and provides various and changing acoustic responses. The goal was to create a space that invites to play with the apparent and hidden affordances found in the room. New experiences and auditory perspectives should be gained during an installation visit, but also connections to real world acoustics should be found. The overall sensation and the character of the installation aimed to evolve a playful environment, where one could change the state of the installation and eventually the whole acoustic interactively by producing sound. For this reason, much time was spent to develop a system, that provides interesting acoustical responses, that either imitate real world acoustics, or create completely new auditory events, not found in nature. The system we designed, works with a microphone array capturing sound and playing it back via 48 loudspeakers.

An aesthetic and conceptual decision was made about limiting the employed signals to sounds produced by an acoustical source at the current moment in the installation space. This implies, no generated or previously recorded audio material was used by the designers in any way for the installation. The installation visitor acts in the surrounding environment and thus creates the input signals at the moment being. Signals are processed, altered and played back by the system. The system response stands in a distinct relation to the action of the visitor, who eventually identifies the connection of action and reaction as a feedback process. Nevertheless, a full understanding of the process is not required. Visitors should be given the opportunity to explore the installation space by walking through the room on random paths. For this reason, as well as to underline the playful approach, all objects were arranged randomly in the room to form clusters. Regular or well-arranged groups were avoided in order to form a homogeneous and organic structure for auditive and visual cognition. Also, the system's behaviour is different for randomly distributed loudspeaker positions, where some effects are more likely to happen or easier to achieve. This topic will be explained in more detail in chapter 3.4.3.

1.2 Staging

As the venue for the sound installation the *Forum Stadtpark* was chosen. *Forum Stadtpark* is located at the center of Graz (Styria, A) in the middle of the urban park. The venue was initiated in the early sixties of the last century and is a place for inter medial art forms with approximately 150 events per year hosting works in the fields of architecture, literature, perfomance, building arts, theater etc. The location has an open design and is flooded with light through the wide glass facade allowing to capture the surrounding park while visiting the installation space.

Interstices - Zwischenräume was opened to the public on 31st August 2013. On this occasion a performance was staged where the installation was used as an instrument. The installation could be visited in the following three weeks. The finale was held on 21st September 2013 enacting a performance/concert showing recent works of Marko Ciciliani, Raphael Kapeller, David Pirrò, Martin Rumori and Gerrit K. Sharma.

The installation is based on previous work of Georgios Marentakis and David Pirrò, who both work at the *Institute of Electronic Music and Acoustics, University of Music and Performing Arts Graz*. *Interstices - Zwischenräume* is founded on the eponymous installation held in ESC Labor in Graz, between 12th and 21st of January 2012 [MP12]. Within the context of his *Toningenieursprojekt*, Raphael Kapeller joined the development team to work on the installation in July 2013. At the same time Johanna Reiner, who works as a visual artist and curator, was asked to design the visual appearance and the scenery of the installation to add a different aspect to the final work.

The fundamental elements of the setup are 22 microphones and 48 loudspeakers forming a feedback system, where the captured sound is played back across the installation space. As alluded before the idea was to act in the installation space and to produce sound. Visitors were encouraged to interact with the system simply by placing sound producing objects in the installation space, hoping that the natural curiosity would lead them to use the objects. The objects included a whistle, a rubber duck, a trampoline with small bells attached to its membrane, a snare drum hanging from the ceiling and a drum stick. 20 microphones formed a regular array hanging from the ceiling at a height of approximately $3m$. The connecting cables were directed towards the middle of the room, forming a starlike web spreading. The remaining 2 microphones were placed outside the installation space at the facade of Forum Stadtpark, one at the entrance and the other at the backside facing the main square of the park. The loudspeakers were attached to objects like wooden panels, stands, etc. and randomly distributed in the room. The cables consisted out of a splaying of 4 single loudspeaker connections which were extended with one multicore towards the amplifiers. All loudspeaker cables (12 multicores, 48 loudspeaker cables) were arranged on the floor forming serpentinous paths.

Further, a vital part of the installation was the artwork of Johanna Reiner. She installed a printer, which printed pictures every five minutes, sent from her smart phone. The paper dropped to the floor, slowly building a heap. A second heap constructed out of shredded paper and a light wooden framework was placed in the installation space. This shredded-paper-heap formed a cave which could be entered by crawling in. 4 of the 48 loudspeakers were placed in the shredded-paper-heap playing back acoustic signals captured by the 2 microphones outside the installation space in the surrounding park.

Figure 1: Overview of the installation space

2 Background

This chapter gives an introduction to topics, theories and tools used later in this work. The focus lies on giving an overview and covering some theory, which is needed for a full understanding of the following chapters. Nevertheless, chapter 3 tries to give full explanations and references this chapter for additional information whenever needed.

2.1 Audio tools

Audio tools are standard applications in audio engineering to shape signals. In a nutshell, the frequency content or the dynamic range of an audio signal can be altered and fine-tuned, in order to obtain a desired outcome.

2.1.1 2nd order peak filter (IIR)

To control the frequency content of a signal a 2nd order peak filter was implemented. A peak filter is a variable tool, where the center frequency (f), the gain factor (g) and the bandwidth (Q) can be separately adjusted. The filter has a *infinite impulse response (IIR)* and is done in *direct form II* [Smi07]. The coefficients and the filter design are taken from [Zöl02]. The filtering is done sample wise applying the recursive difference equation (1) in the time domain which allows a fast performance.

$$y(n) = b_0 \cdot x(n) + b_1 \cdot x(n-1) + b_2 \cdot x(n-2) - a_1 \cdot x(n-1) - a_2 \cdot x(n-2) \quad (1)$$

The coefficients a_1, a_2, b_0, b_1, b_2 have to be calculated in advance, because of the dependency of the three parameters f, g, Q. The algorithm was tested and implemented in *Matlab* before it finally was set up in *rattle*. This simple equalizer consists of one full parametric band, which was sufficient for the application. Nevertheless, implementing further bands when needed is an easy task. The 2nd order peak filter is used in all outputs of the system.

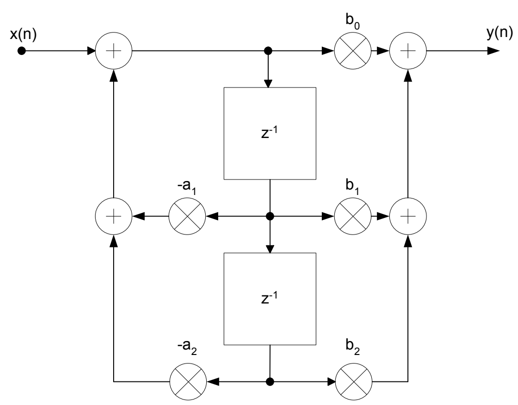


Figure 2: Block diagram of the 2nd order peak filter (IIR) in *direct form II*

Table 1 shows the calculation formulas for the coefficients for positive (boost) and negative (cut) gains G, where $K = \tan\left(\frac{\pi f_c}{f_s}\right)$ [1].

	boost $V_0 = 10^{\frac{G}{20}}$	cut $V_0 = 10^{-\frac{G}{20}}$
a_1	$\frac{2\cdot(K^2-1)}{1+\frac{1}{Q}\cdot K+K^2}$	$\frac{2\cdot(K^2-1)}{1+\frac{V_0}{Q}\cdot K+K^2}$
a_2	$\frac{1-\frac{1}{Q}\cdot K+K^2}{1+\frac{1}{Q}\cdot K+K^2}$	$\frac{1-\frac{1}{Q}\cdot K+K^2}{1+\frac{V_0}{Q}\cdot K+K^2}$
b_0	$\frac{1+\frac{V_0}{Q}\cdot K+K^2}{1+\frac{1}{Q}\cdot K+K^2}$	$\frac{1+\frac{1}{Q}\cdot K+K^2}{1+\frac{V_0}{Q}\cdot K+K^2}$
b_1	$\frac{2\cdot(K^2-1)}{1+\frac{1}{Q}\cdot K+K^2}$	$\frac{2\cdot(K^2-1)}{1+\frac{V_0}{Q}\cdot K+K^2}$
b_2	$\frac{1-\frac{V_0}{Q}\cdot K+K^2}{1+\frac{1}{Q}\cdot K+K^2}$	$\frac{1-\frac{1}{Q}\cdot K+K^2}{1+\frac{V_0}{Q}\cdot K+K^2}$

Table 1: Coefficients for the 2nd order peak filter [Zöl02]

The filter algorithm was tested in *MATLAB*, but when implementing it in *rattle* problems occurred during the testing. With low center frequencies (nearby 0Hz) the algorithm was not stable. The reason for this behaviour was found in the appearance of denormals. Compare with chapter2.2. Solving this issue was done with the later described algorithm.

2.1.2 Compressor

A compressor is used to control the dynamics of a signal. It is used to dampen signal parts with higher amplitudes. This gain reduction is compensated by a makeup gain, that amplifies the whole signal. Hence, the ratio between signal parts with lower and higher amplitudes is changed and lower amplitudes are perceived louder.

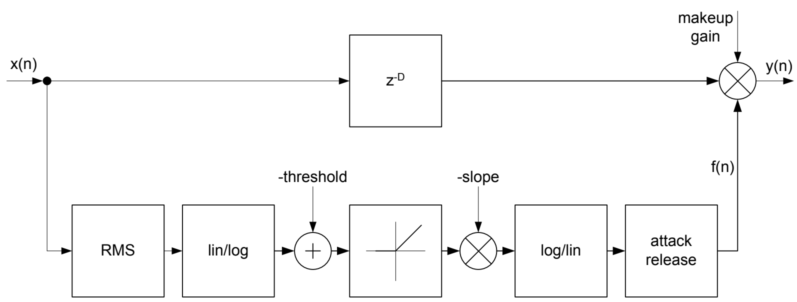


Figure 3: Block diagram of the implemented compressor [Zöl02]

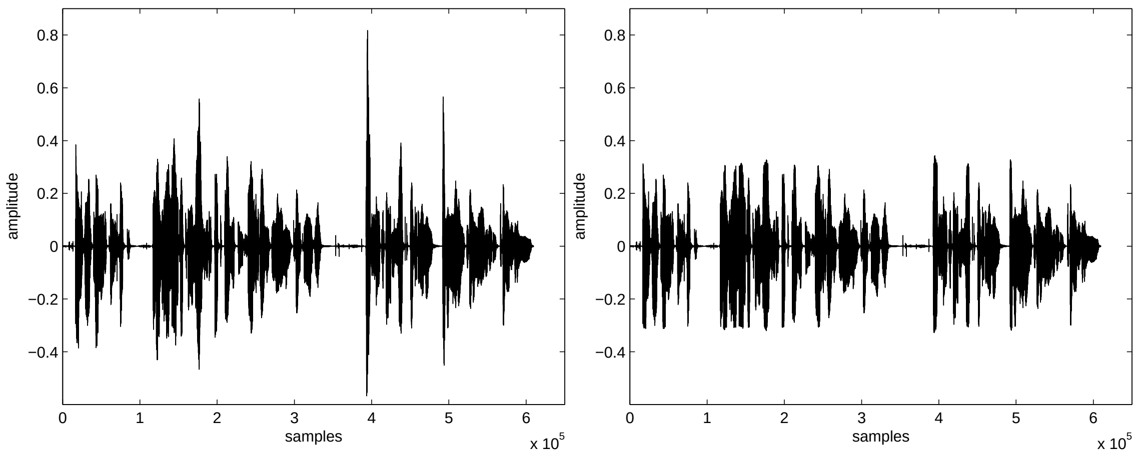
1. f_c ... center frequency; f_s ... sampling frequency

Figure 3 shows the block diagram of the implemented compressor. The absolute value [2] of every sample of the input signal $x(n)$ is calculated. If the RMS value of the inputsignal $x(n)$ is bigger than the defined threshold, the signal part above the threshold is multiplied with a slope S, to achieve the desired ratio R.

$$S = 1 - \frac{1}{R} \tag{2}$$

The attack and release time define the characteristics and the behaviour of the compressor. As the default values the attack time is set at $0s$ and the release time is set at $1s$, which cuts all percussive signals and allows to louden the signal by amplifying the compressed signal with the makeup gain.

Two advantages are given by changing to the logarithmic domain (i) mathematical: multiplications in the linear domain are additions in the logarithmic domain [3] and (ii) psychoacoustic: the human ear follows the logarithmic scale for perceiving loudness [Zöl02].The implementation of the compressor is done using two exceptional cases. The first is done by determining if the threshold has been passed and compression has to be made $if\ x > threshold$. Following this distinction without any transition area is defined as *hard knee*. The second exceptional case is $if\ a < 0$ where $a = x_n - x_{n-1}$ and defines if the current sample has a higher value as the previous one and therefore the attack time has to be taken into account. On the other hand, $if\ a > 0$ the release time is used, because the signal already reached it's maximum and is decreasing. Figure 4a shows a speech signal with peaks in the amplitude and Figure 4b shows how the implemented compressor dampens these peaks beyond a threshold of 0.3. In order to see the changes quite easily no makeup gain was applied in the graph.



(a) speech uncompressed (b) speech compressed

Figure 4: Function of the compressor for a speech signal. $threshold = 0.3$, $R = 8$, $release = 0.01s$, $attack = 1s$, $makeup = 0$

2. For one sample the absolute value and the RMS value are equal. $\mid x_1 \mid = \frac{\sqrt{x_1^2}}{1}$
3. Of course the calculation of the logarithmic values also costs CPU-performance.

2.1.3 Signal amplitude fading

Signal amplitude fading is used to avoid audible *clicks* while playback. Clicks occur when the audio signal waveform shows discontinuities. This means, that the waveform jumps from one value to the next discretely. For example, this happens when the playback does not start at a zero-crossing, forcing the output to jump from zero to the desired value of the playback. Another example is given by jumping between signals without crossfading them.

Signal amplitude fading is done with an integrator, that multiplies the signal with a envelope in form of an *Euler's function*. Therefore, hard switching or jumping between signals is prevented. The algorithm checks, if the incomming signal is different to the previous one and applies the fade if the absolute value varies for more than $-60dB$ (± 0.000001). The fade is done by adding the weighted signal difference ($newsignalvalue - oldsignalvalue$) to the old signal. The weighting is done by the Euler's function $e^{\frac{log(0.001)}{s \cdot fs}}$.

Pseudo code for signal amplitude fading: [4]

```
  sig=[sig[0],....,sig[n]];
  fac = exp(log10(0.001)/(sec*44100));
  res = 0.0;
  i=0;
  ao = 0.0,
  out = 0.0;

for i=[0,....,n]{
  if ( res - sig[i]) > 0.000001) {
        ao = sig[i];
        res = ao + (fac*(res - ao));
      }
  else {
        res = aIn;
      }
end
```

2.2 Denormals

Denormals are also called subnormals and refer to numbers closest to zero in the floating point representation. The floating point format consists of a sign bit, exponent bits and mantissa bits. In the case of IEEE 754 single-precision binary 32bit floating point format these bits are 1 sign bit, 8 exponent bits and 23 mantissa bits. [dS05] Denormals have the smallest possible exponent and the most significant bit (MSB) is zero [DK06]. Because of the small value of denormals the process gets slowed down and the performance impact is huge. The simplest way to solve the problem of appearing denormals is to define

4. this example shows the functionality of the algorithm without using a correct syntax.

an exceptional case, introducing an value offset, thus defining the smallest possible value. Although applying exceptional cases costs computational power, the gain of computational power by eliminating denormals exceeds.

```
float zapgremlins(float x)
{
    float absx = fabs(x);
    return (absx > (float)1e-30 && absx < (float)1e30) ? x : (float)
        0.;
}
```

In this case the absolute value of x is set to zero if it is smaller than $10^{-30} \simeq 2^{-100}$. With the IEEE 754 single-precision binary 32bit floating point format the smallest number representable by the exponent is $2^{-2^8} = 2^{-128} \simeq 3.4 \cdot 10^{-38}$. Hence, the precision of the calculation is set down in the exponent from 10^{-38} to 10^{-30}. There is also a upper boundary for high values at 10^{30} where the value is set to zero as well when x is passing this threshold. Setting the value to zero prevents it to grow bigger than the highest possible represented number. This occurs when all 32bits of a single-precision floating point format are set to 1 but yet another increase is added. In this case the value is no longer a number but set to infinity inf. The introduced function is only capable of stopping enlarging processes, once inf is reached [5] the function is futile.

2.3 Physical modelling theory

This chapter summarizes very shortly two paradigms explaining physical modelling. The final installation uses physical modelling to control the current state of the system. For a detailed explanation of this utilization, please refer to chapter 3.5.1.

2.3.1 Models and their appropriation

A model is used to depict a process, an event, a scenario or an application found in nature or in virtual space. The model enables a better understanding of the course of events and allows to show either an overview or a zoomed in point of view. Problems can be seen more easily and simulations help to find possible solutions. Also, a model uses always simplifications of its world. On the one hand this is done to separate problems and show ideal behaviour of systems and on the other hand in most cases it is not possible to model or simulate every aspect found in the system's world.

2.3.2 Physical modelling

Physical modelling is used to model real world dynamics as well as dynamics not found in nature [Hen04]. As the name implies, the modelling uses elements defined and described in physics. Differential equations using these elements can be set up and solved by

5. For example directly by $\frac{x}{0} = inf$

standard solving algorithms. This means, that dynamic problems are explained by solving the differential equations of the model. The two elementary objects are *mass* and *link*. Decisions have to be made, which objects are represented by masses and how the interaction works between them . The relationship between masses is defined by the link and is formulated as a force acting on the masses. Different force types can be used (spring like, gravitational, etc.). Therefore, *Newtonian dynamics* can be used to put mass and link into one fundamental equation:

$$F = m \cdot a \tag{3}$$

The second equation is defined by the force type. Here, the link is a visco-elastic connection between two masses [Hen04]

$$F = k \cdot x + d \cdot v \tag{4}$$

object	mass	distance	time	velocity	acceleration	gravitational constant	stiffness	damping
symbol	m	x	t	v	a	G	k	d

Table 2: Objects and their symbols.

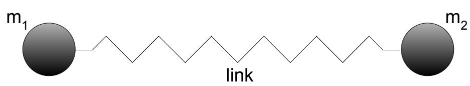


Figure 5: Mass and link

Equation 4 follows *Hook's law* for spring like forces. The second part of the equation brings damping into account. Another force type is given by *Newton's law of universal gravitation*, where the force between two bodies is directly proportional to the product of their masses. [gra]

$$F = G \cdot \frac{m_1 \cdot m_2}{x^2} \tag{5}$$

Equation (4) and (5) can be rewritten to form a global equation, where the values of the variable α applied in the formula define, if the force is spring-like or gravitational.

$$F = k \cdot x^{\alpha} + d \cdot v \qquad \text{with } \alpha = [-2; 1] \qquad \text{and} \tag{6}$$

$$k = G \cdot m_1 \cdot m_2 \tag{7}$$

In this case m_1, m_2 and k are fixed by parameters, hence $G = \frac{k}{m_1 \cdot m_2}$. Using equation (3) and equation (6) with a damping of $d = 0$ leads to

$$F = \quad k \cdot x^{\alpha} = \quad m \cdot a \tag{8}$$

Solving this equation for the acceleration a leads to

$$a = \quad \frac{k \cdot x^{\alpha}}{m_1}$$

Also, the acceleration is the second derivation of time. This two expressions can be equated

$$a = \quad \frac{d^2x(t)}{dt^2} = \quad \frac{k \cdot x^{\alpha}}{m_1}$$

Putting a back into equation (6), the introduced forces can be expressed by ordinary differential equations (ODE) of second order $F = m \cdot \frac{d^2x(t)}{dt^2}$. The second order ODE can be rewritten as first order equations, where two iterations are needed to form the second order differential equations.

$$v = \quad \frac{dx(t)}{dt} \tag{9}$$

$$\frac{dv(t)}{dt} = \quad \frac{F(t)}{m} \tag{10}$$

For the simple case of two masses, an analytical solution of the first order ODEs is possible, but for more than two interacting particles it can be impossible or at least very difficult. Therefore, numerical approximation has to be used. The idea behind the approximation is to discretise time (timesteps $\Delta t = \frac{1}{f_s}$) and to replace the DE by a approximated version. The simplest and fastest way to do this is *Euler's method*, where the linear approximation is used to find the next value of the movement in time along a direction x. Here, the next value is calculated by slope times timestep plus the old value $x_{n+1} = x_n + \Delta t \cdot \frac{dx(t)}{dt}$ (which is the same as equation (16)) . The sequence of the calculation steps is important. For the discrete, numerical solution the differential d changes to Δ, in order to show the difference between continuous and discrete solving approaches. Here, one can see the Euler's method to calculate x with two iterations of solving differential equations.

$$\frac{\Delta v(t)}{\Delta t} = \quad a \tag{11}$$

$$\Delta v(t) = \quad a \cdot \Delta t \tag{12}$$

$$v_{n+1} = \quad v_n + \Delta v(t) \tag{13}$$

$$v_{n+1} = \quad \frac{\Delta x(t)}{\Delta t} \tag{14}$$

$$\Delta x(t) = \quad v_{n+1} \cdot \Delta t \tag{15}$$

$$x_{n+1} = \quad x_n + \Delta x(t) \tag{16}$$

For the calculation, initial values for exerted force are given. Also the parameters k and m are known and therefore the movement due to the force can be calculated. Nevertheless, a changing position of a mass introduces another force formed by the connection of the link. Depending of the force type defined by the link (gravitational, spring like), the force reacts as a counter force pushing the connected masses in certain directions. Hence, without attrition a perpetuation oscillation is the result, where movement and forces are calculated in dependence of each other.

The main disadvantage of the Euler's method is it's inaccuracy as well as the possibility to become unstable [Zhi]. Nevertheless, the algorithm is used in *rattle*, and works well as long as objects with high masses are used. In this case the inaccuracy is small, because with low velocity / low acceleration the changes over time are small as well and the approximation using only the slope works with sufficient precision for the time being.

In the future Euler's method will be changed to use the *Verlet algorithm*. The Verlet algorithm is more precise and stable. The error produced by Euler's method is of second order $O(\Delta t^2)$ versus the error of the Verlet Algorithm is $O(\Delta t^4)$.

2.3.3 Fields

In the previous chapter the influence of forces exerted onto connected masses and their movement has been explained. Another way to take a look at the law of movement is to define fields. *A field is a physical quantity that has a value for each point in space and time.* [Gri98] A mass forms a field and changes its surrounding space. If there is a second mass, a force is exerted onto this mass depending on the field intensity [NS14]. This means the field intensity A can be described as following:

$$A = a = \frac{F}{m_2} \tag{17}$$

$$A = -G\frac{m_1}{r^2} \qquad \text{for gravitational fields} \tag{18}$$

It can easily be seen, that the field paradigm uses only one mass at the center of the field, whereas a second mass is not present at the definition of the field strength (m_2 cancels in equation (18)). As a consequence, every mass is surrounded by its own field. The field attributes are determined by the physical equations derived from the force types described in chapter 2.3. The field extent depends on the parameters shown in table 2. In other words, the field paradigm allows a global definition of the physical relationships caused by a mass where the linking attributes are taken into account. The field defines a calculation law where forces can be determined for every point in space and time. This is the main difference to the force paradigm, where the outcome is not a law, but already a force specified for a certain point in space.

2.3.4 Physical modelling using *rattle*

***rattle** is an efficient implementation of a mass-based physical modelling server, written in C.* [MP12] Simulation and prototyping of a model behaviour can be achieved rapidly.

The model is constructed using particles (masses) and connecting them to each other. In rattle the field paradigm is used, where characteristic values like mass, attrition or force type are attached to every particle. These characteristic values define the field, the particle and the interaction between them. Forces can be exerted onto masses and *locations , displacement, velocities, acceleration and energies of the elements of a model are sampled at variable rates.* [MP12] With variable sample rates, rattle can be used either for sound generation (sample rate at standard audio rates) or for controlling a modelled object or application (sub-audio rates) by changing parameters with the fluctuating locations, displacements, etc. introduced by the simulation. Real-time parametrization is featured, where modifications are done without having to re-compile the whole code.

2.3.5 Applied fields

For the installation three different fields were applied. For a detailed utilisation refer to chapters 3.5 and 3.5.1. Here, the concepts for the implemented fields will be covered. All fields used in this context work similar, but the detail makes the difference.

Field type 1: Field type 1 is used to keep a mass at its position by using field attributes. In simple words, the mass is attached to a fixed position, but movement around the equilibrium is allowed. When the mass is forced to move away from the equilibrium position by another field, hence a force introduced by a second mass, the distance between the fixed position and the current position is calculated. The distance determines the force strength and increases proportional to the distance for a spring like force. As a consequence, the increasing force pulls the mass back towards its stable location.

Field type 2: Field type 2 works between 2 masses. This is slightly more complex, as the masses, and therefore also both fields, have to be taken into account because of their interaction with each other. Also here, the distance between two *linked* masses are calculated and the resulting forces are summed. As a consequence, the net force moves the masses, which leads to a new iteration of the calculation.

If a regular array of masses with fixed positions and fields of both types is used, a special case emerges. Problems occur at the borders of the regular array because of the disparity of force for the elements at the border. One solution is to use *helper masses*, which surround the array. The helper masses are fixed at their position. The helper mass field exerts forces onto their neighbouring masses to compensate the disparity of force. Helper masses do not take part in any other application. For example exerting a force onto a helper mass, has no effect and the force vanishes by definition.

2.4 Localization

For many applications of the desired outcomes of the installation, it is beneficial to know where the sound captured by the microphones has it's origin. Localizing the sound source position enables multiple possibilities for using the data. For example the position can be used to decide where the signal will be represented by the loudspeakers (nearest or most faraway position) or the installation space could be divided into certain zones with different behaviour of the system. Therefore a simple and efficient tool was desired to localize sound. In the following section two models of localization algorithms will be described.

2.4.1 Localization using beamforming

One idea to get information about the sound source position was to use the microphone array and to implement a *beamformer*. This beamforming algorithm is based on the time delay of arrival (TDOA) measurements for every microphone in the array. This means, depending on the sound source position, soundwaves arrive at different times at the microphones: The differences between arrival times are called *time differences* τ_{ij}. *Range differences* d_{ij} are defined as the distance difference between sound source and the microphones i and j. In order to determine the exact position of the source in space, the position of the microphones, the speed of sound (depending of the air temperature) and the time difference τ_{ij} for every microphone pair has to be known. [HBEM01]

$$R_i = \| r_i \| = \sqrt{x_i^2 + y_i^2 + z_i^2}, \quad i = 1, ..., N \quad \text{microphone position (known)} \tag{19}$$

$$R_s = \| r_s \| = \sqrt{x_s^2 + y_s^2 + z_s^2} \quad \text{source position (desired)} \tag{20}$$

$$D_i = \| r_i - r_s \| = \sqrt{(x_i - x_s)^2 + (y_i - y_s)^2 + (z_i - z_s)^2} \tag{21}$$

$$d_{ij} = D_i - D_j = c * \tau_{ij} \tag{22}$$

$$\tag{23}$$

τ_{ij} can be determined by calculating the cross-correlation between the signals of every microphone pair ij. [VMRL03]

$$C_{ij}(\tau) = \sum_{n=0}^{N-1} x_i[n] x_j[n - \tau] \tag{24}$$

$C_{ij}(\tau)$ is at it's maximum when the two signals x_i and x_j are matching; in other words when $\tau = \tau_{ij}$. Equation (24) shows the convolution of two time signals. By converting the signals to the frequency domain, the convolution in (24) becomes a multiplication: $X_i(k)X_j^*(k)$. Inverse fourier transformation (IFFT) provides the correlation in the time domain.

$$C_{ij}(\tau) = \sum_{n=0}^{N-1} X_i(k) X_j^*(k) e^{\frac{i2\pi k\tau}{N}} \tag{25}$$

Using the FFT and IFFT algorithms is done to fasten the process and to lessen the computational cost. Performing the convolution in the time domain would not be as efficient as the fourier transformation.

The correlation of signals of neighbouring microphones have to be calculated. Figure 6 shows the microphone array. Four microphones form a square that is divided into four identical quadrants. The length $a_1, a_2, b_1, ..., d_1, d_2$ define the borders of the quadrants and the range for the time difference τ_{ij} of the crosscorrelation. For all 22 microphone pairs the crosscorrelation in all 4 quadrants of all 9 squares is calculated. Theoretically the quadrant containing the sound source would provide a maximum. One limitation of this method is that most signals have a low pass characteristic and result after crosscorrelation in broad peaks. This can be improved by whitening the signal, which is done by normalizing the convolution in the frequency domain $\frac{X_i(k)X_j^*(k)}{|X_i(k)||X_j^*(k)|}$ [VMRL03].

Using the beamformer, one has to decide how the microphone array is set up. The aperture of the array defines the valid frequency range in which the localization delivers robust solutions. On the one hand small distances d between neighbouring microphones in the array avoid *grating lobes* for increasing frequencies. Grating lobes are mirrored main lobes (spatial aliasing) and a special case of side lobes, because of their high amplitude [Wik14]. On the other hand distance $L = N \cdot d$ between the furthest apart microphones should be large for a small beamwidth and a good resolution [Per14] . This two requirements can be achieved, but only with high expenses.

Thinking of a limited number of microphones, in our case two arrangements for the array were possible:

- small array with small distances between microphones: This means $L <$ and $d <$, which gives a good angular resolution, but no information about the distance.
- big array with big distances between microphones: This means $L >$ and $d >$, which has a good resolution, but the upper frequency limit to avoid spatial aliasing is lower.

As a conceptual decision we already had chosen a wide regular array distributed in the room. Knowing the restrictions by the model, nevertheless we gave it a go and tried to implement this algorithm in the installation. Testing this application provided non-satisfying results. It was not possible to find a unique solution for the position of the sound source. Another problem is that the sound source is within a distance of the array that is smaller or at least similar to the array aperture. This means the far-field constraint for beamformers [Hof08] is violated because the signal can not be assumed as a plane wave. Another problem emerges because of spatial aliasing. The upper frequency limit for a microphone array can be calculated by:

$$f_{max} = \frac{c}{2 \cdot \Delta x \cdot \sin(\alpha_{max})} \tag{26}$$

$$c \approx 331 + 0.610 \cdot t_{air} \tag{27}$$

With $\alpha_{max} = 90°$ and $\Delta x = 2m$ the upper frequency limit is only at $f_{max} = 86.6Hz$.

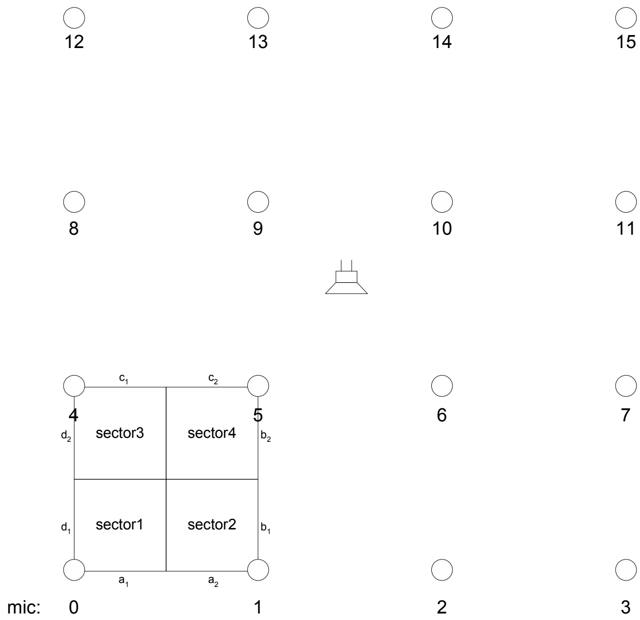


Figure 6: Microphonearray seen from above.

2.4.2 Localization using a simple root mean square (RMS) algorithm

A RMS algorithm is the simplest form to calculate the average value of a signal and is done sample-wise by equation (28):

$$\bar{x}_{RMS} = \sqrt{\frac{x_0^2 + x_1^2 + x_2^2 + \ldots + x_{n-1}^2}{n}} \tag{28}$$

The RMS value is calculated for all 16 microphones for blocks of 1024 samples and stored in an array. The *hopsize* is set to 256 samples, what means that every 256 samples a new RMS value is calculates over 1024 samples.This simple algorithm allows a robust localization of the sound source by calculating it's RMS power. The microphone with the highest RMS power input defines the position of the sound source. Of course, only the discrete positions of the microphones in the room can be detected. An improvement of this algorithm was achieved by using a *lag*, which smooths the discrete positions to a continuous function. Also a physical modelling of the sound source was tested, but finally not implemented, because of the good results and the lower computational cost of the *lag*-function.

Pseudo code of RMS-algorithm:

```
i = 0;
counter=0;
size=1024;
hopsize=256;
```

```
sig=[sig[0], sig[1], .... sig[n]]

if (counter >= hopsize){
   rms = 0.0;

   for i=[0,....,size]{
        squared_signal= Sum(sig[i]*sig[i]);
        rms = Squareroot(squared_signal)/size;
        counter = 0;
       }
else{
      counter = counter+1;
       }
end
```

3 Designing Process and Implementation

In the following section the process of implementing and adjusting the sound installation from the very beginning will be documented and described. Especially the decisions on the way to the final product will be highlighted and discussed considering it's technical and aesthetic aspects. The process was following an exploratory approach, where the way to realize the basic design ideas where done step by step and iteratively to develop the final version of the installation.

3.1 Equipment

Interstices works with 48 loudspeakers, 22 microphones, 3 eight channel preamps, 6 eight channel AD/DA converters, 1 madi/adat converter and a PC running on *Linux*. Figure 7 shows the signal flow through the system as well as the used hardware components. Table 3 gives an overview of the used products and a brief description of their applications.

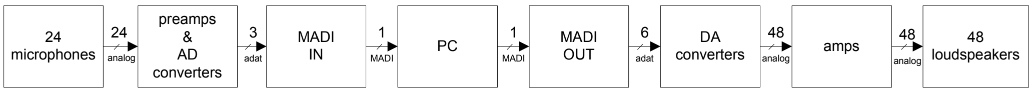


Figure 7: Blockdiagram of the hardware components and the signal flow of the system.

item	product	amount	description
microphones	Behringer ECM8000	22	measurement condenser microphone
preamps & AD converters	Presonus DigiMax LT	3	8 channel digital mic/line preamp with *adat* out
MADI converter	RME ADI-648	1	format converter from MADI to *adat* and vice-versa
DA converters	RME ADI-8 DS	6	8 channel AD/DA converter
amps	IEM custom made	1	class d amps for 48 loudspeakers
loudspeaker	IEM custom made	48	loudspeakers

Table 3: Equipment.

For the final installation some editing was made to the existing set up of the equipment in order to gain maximum flexibility, as well as reducing the time needed for the assembling. Therefore the outputs of the amps where changed from *banana plugs* to 8-pole *speakOn* connections. Using 8-pole connectors 4 loudspeakers can be gathered together by a splaying and extended with one cable. Also, all former described 19-inch devices were put together in one rack, to finally result with two racks (amping, side rack).

3.2 Setup

The following chapter describes the setups used for testing in the *CUBE* of the *Institut für Elektronische Musik und Akustik IEM Graz* as well as the setup for the final installation in *Forum Stadtpark*.

3.2.1 *CUBE* setup

In preparation of the final installation set up in *Forum Stadtpark*, the hardware and sofware needed for the project was developed and tested in a prelimiary phase during approximately 4 weeks in July and August 2013. For this porpouse a smaller setup (16 microphones, 48 loudspeakers) was used at *CUBE* of the *Institut für Elektronische Musik und Akustik IEM Graz*.

Similar to the installation in *Forum Stadtpark* the microphones were hanging from the ceiling at a height of approximately 2.20m forming an regular array of 4x4 microphones. Objects like tables, chairs, stands or wooden panels where distributed in the room forming clusters. The objects served as placing positions for the 48 loudspeakers, which were also randomly distributed at different heights. In the middle of the room the side racks containing preamps, AD/DA and MADI converters and amps were placed. From there the microphone cables were strained from the side rack directly to the ceiling and split to the different microphone positions. The next step was to determine the positions of all microphones and loudspeakers. As the microphones form a regular array, the coordinates of the loudspeakers were measured as a function of the microphone positions.

Figure 8: Overview of the sound installation at *CUBE*

Figure 8 shows the installation in *CUBE* before the editing of the hardware. The gray loudspeakers are connected directly to the amping with cables of about 3m in length. Clusters are formed, but the possible positions of the loudspeakers are restricted by the cable length. A fundamental decision was to arrange the loudspeakers in a none regular way. This was done for aesthetic reasons as well as to enable the possibility to form clusters. Also, walking through the installation should be able on none regular paths, designed by the arrangement. Particularly with regard to auditive experience in the installation a none regular arrangement is beneficial, because various outcomes (e.g. more resonances) are possible.

The microphones plus preamps were calibrated using a pistonphone at 94dBSPL $@1kHz$. The measurement was done in *Pure Data* (PD) by tuning the input signal via the gain control of the preamp to 88dB for all microphones. At a later time during the process the gains where changed to a lower value (76dB) to achieve a less sensitive feedback system.

The loudspeakers, amps and DA converters were tested using white noise produced in PD and played back consecutively by all loudspeakers. Later also the loudspeaker outputs were calibrated or flattened according to the room. (Compare with chapter 3.6)

3.2.2 *Forum Stadtpark* setup

As mentioned before, the setup in *Forum Stadtpark* is the further development and enhancement of the tested *Cube* setup. The main differences are the increased number of used microphones, the additional usage of a subwoofer and of course the enlargement of the installation space.

Figure 9 shows a picture of the final installation. Likewise to the testing phase, the microphones are forming a regular array hanging from the ceiling at a height of $3m$. Yet, the array consists of two rectangular parts, where the first part is build of 16 microphones, whereas the second part consists of only 4 microphones. The T-shaped array follows the room layout, in order to cover the whole installation space. The loudspeakers are placed irregularly on different objects randomly distributed in the room. The objects include chairs, stands, wooden panels, a trampoline and the previously described heap of paper. Another enhancement, based on an artistic and aesthetic decision, is utilising 4 microphones placed outside the installation space, at the entrance and at the backside of *Forum Stadtpark*. These signals are directly routed to the four loudspeakers in the shredded-paper-heap. Recapitulating, 20 microphones and 44 loudspeakers are used for the main part of the installation and another 2 microphones and 4 loudspeakers are used to play back environmental sounds in the shredded-paper-heap. This leads to a total number of 22 microphones and 48 loudspeakers.

Figure 9: Overview of the sound installation at *Forum Stadtpark*

In the following chapters, the process of developing the installation will be described. Most of the applications and functions were implemented and tested during the preliminary phase in *CUBE* and carried over to the final version in *Forum Stadtpark*. To avoid confusion and to be consistent, the final version as implemented in *Forum Stadtpark* with 22 microphones and 48 loudspeakers will be used for further explanations.

3.3 Microphone to loudspeaker signal routing

This section will describe how the patching of the inputs to the outputs of the system is made. The emerging question is how one can distribute the signals of 20 microphones to 40 loudspeakers. One possible solution is to use the positions of the microphones and loudspeakers to decide how the assignment works.

The positions of the microphones in the array are known. In the model the array is centered around the origin in the $xy-$plane and elevated to the measured hight of the hanging microphones. As mentioned before, the coordinates for the loudspeaker positions are measured as a function in the grid of the microphone array and the coordinates are easily calculated. Different methods were tested to produce an interesting routing, that supports the goals of the installation. For example like changing the perceived room size.

The first idea was to map every microphone signal to the nearest loudspeakers in space. Calculating the nearest loudspeaker is done by a simple vector function where for 20 microphones all 48 distances to the loudspeakers are calculated. This results in an *array* of $20 * 48 = 860$ values where the first 48 values correspond to microphone 1, the next 48 to microphone 2 and so on. Finding the microphone nearest to a loudspeaker is done by searching for the minimum value in the array and dividing the array-number by 20 to find index i. The *modulo-function* provides the desired loudspeaker index j. This is done by multiplying the rest of the devision with it's divisor, hence the inverse calculation step is performed to restore a natural number j. The indices i, j are stored separately. Once this task is accomplished the minimum value is set to a high value. The aforementioned steps are performed again to detect the currently smallest value.

Restrictions are made where one microphone only delivers signals to 2 loudspeakers and no repetition of loudspeakers is allowed. This means, only one microphone signal is played by one loudspeaker, but two loudspeakers get the same signal, hence every loudspeaker is playing. This kind of distribution introduces with it's restriction some problems. At the last iterations of the process the left over loudspeakers are likely to be far apart and the signal is not passed to the nearest loudspeakers, but to the only left ones. In this case the spatial coherence of sound source captured by the microphones and position of the loudspeaker representation was not given and we decided to find a better solution to maintain the coherence. Adapting the algorithm and making no constraints about the number of loudspeakers per microphone gives a better distribution of the signals to near loudspeakers, but here problems occur because of the unequal number of loudspeakers connected to one microphone. This leads to an asymmetric distribution of sound in the loudspeaker array.

At this point the two tested solution with and without restrictions were kept for future reference and no final decision was made. In chapter 3.4.3 this topic will be discussed in more detail for the distribution of signals sent from one loudspeaker to another.

Alternative microphone to loudspeaker signal routings: Following two routings where also tested, but finally not implemented:

- biggest distance of microphone to loudspeakers: For this routing the spatial coherence of source to representation gets lost.
- zones: In this case the room was divided into different zones where every zone used one of the three scenarios described in the next chapter. This was done because the merging of the scenarios was found to be quite difficult. Nevertheless, the organic behaviour of the installation gets lost due to the introduction of zones and therefore it was not used in the installation.

3.4 Scenarios

The acoustical signals captured by the microphones are reproduced by the loudspeakers. Because of the huge variety of possible approaches to perform this task, the outcome can be tuned into certain directions. For us, the focus lay on altering the spatial perception of the room and composing strategies to change the room. During the testing and parametrization phase we could develop a huge repertoire of different audible system outcomes. The goal was to create interesting sounds, that are connected to certain room attributes, as well as new sounds not found in nature. Because of the exploratory nature of investigating and producing the installation, the process offered new and unexpected impressions. Here, the designer has to make aesthetic decision about a further development and utilization of certain system outcomes. The different explorations can be grouped into categories, which we called scenarios. This was done to give them a distinct nomenclature.

3.4.1 Feedback scenario

Considering the arrangement of the microphones and loudspeakers and how signals are routed, the structure of the system forms a feedback loop. Sound captured by microphones, is played back by the loudspeakers, which again is captured and played back. Therefore the closed feedback loop will form different loop gain factors for all frequencies, depending on the system's attributes. The highest loop gain factors are called resonances which will be audible as feedbacks. Feedbacks[6] have a sinusoid character and occur, as described above, when the output is feed back and added to the input. The electronic as well as the acoustic parts determine the feedback frequency [Wei08]. For the acoustic part the room attributes like the size, the reverberation time or the surface materials used in the room are responsible for the feedback frequency. On the electronic part every single device in the system contributes in the transfer function of the electronic system. The gain factor parameter is of special interest, because it sets the threshold when acoustic feedback occurs without changing the system, which defines the feedback attributes. At hardware level the gain factor can either be influenced by the level of the microphone-preamping or by amplification of the output signals.

Immediately after putting up the setup and connecting the inputs to the outputs, depending on the gain factor of the system, feedback is audible. Without changing any attribute in the system, the highest loop gain (resonance) is audible at a fixed frequency. This time invariant behaviour for static systems is of great benefit for destroying feedbacks, as it is done for example for live performances or concerts with monitoring systems for the artists. For the installation we wanted to work with feedback and not suppress it. Also, the feedback should change by interacting with the system. The question is how one can achieve this task. It is crucial to balance the gain factor nearly to the threshold after which feedback occurs, but stay underneath this border. Then the system is at a unstable balance where small changes in any transfer function of the subsystems are noticeable.

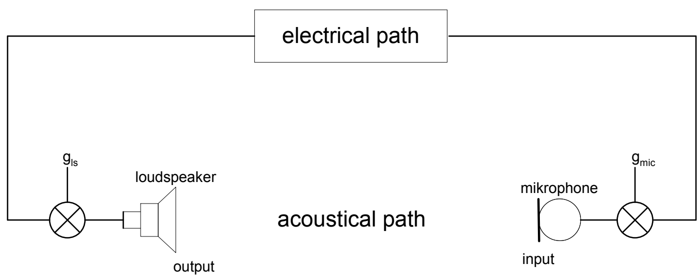


Figure 10: Feedback loop system

In this scenario we decided to work on the acoustical part of the system. Promising results

6. In this paper the term feedback is always used for acoustical or audio feedback as described in this paragraph.

were achieved by entering the installation space or covering one or two loudspeaker-membranes with ones hand. Immediately the acoustical part of the system is changed, other resonances are introduced and because of the unstable balance of the system, in many cases the altered loop gains produce audible feedbacks. Nevertheless, for safety reasons, a limiter function has to be built into the output channels in the software to prevent the signals to overshoot the amplifier inputs and of course to limit the maximum playback volume of the installation to a comfortable level. The parameters of the limiter highly effect the audible feedback. The attack as well as the release time have to be tuned carefully (together with the gain factor) to avoid a too slow reaction to the input signals and effects like *ducking*[7]. The parameters were chosen to support the appearance of feedbacks but to limit their maximum played back volume. With a high thresholds of the limiter, the feedback is likely to stay on a static position, where no changes in the acoustical part show any effects anymore. Therefore, the situation of unstable balance is the most effective status for interacting with the system and the threshold should be chosen to support this situation. Another advantage with lower thresholds is, that also small noises like steps or whispering are clearly audible when played back by the system.

3.4.2 Delay scenario

The position of every microphone and every speaker is known and the routing is fixed. Hence, the distances between source and sink are known as well, and it is an easy task to calculate the time, sound requires to cover the distance. In the following this time will be referred as delay and is calculated by:

$$t = \frac{distance}{c} \cdot f_s \qquad \text{[samples]} \tag{29}$$

By introducing small delays (50 - 22050 samples) in the outputs of every signal feed to the speakers, the whole system changes. Once the routing is fixed all loudspeakers play the signals immediately and together at the same time. Nevertheless, the delay resulting of the distance between sound source and microphone position, as well as different levels in the amplitude and anisotropic frequency responses of the captured signals are automatically part of all scenarios. The feedback scenario depends on the gain factor and the threshold of the limiter. As described above this is the main parameter to find the instable balance of the system. In the delay scenario the most important parameter is the *delay factor* q_{dfac}. The delays are different for every loudspeaker and correspond to the distance between microphone and assigned loudspeaker. Therefore, the time dependent sound wave propagation is taken into account, which means, that the played back sound is delayed with the time sound would need to cover the distance. The following paragraph will describe the function of the buffer system used in the implementation of the delays in *rattle*.

7. ducking: a signal is attenuated by another signal. e.g: commonly used by DJs when speaking while music is playing. In this case the music is turned down automatically. [Izh10]

Buffer system: Every microphone signal is stored in a buffer with the size of several seconds. Also, every loudspeaker has it's own buffer (buffersize= $30*44100$; 30 seconds), where the loudspeaker objects write the signals read from the microphone buffer. For playing back the signal the stored data is read from a delayed position of the loudspeaker buffers. The delays correspond to the distances between microphones and loudspeakers. In the final version signals from other speakers, depending on the loudspeaker mapping and the physical modelling, are added to the output with separate delays. Consequently, the reading point is behind the writing point of the incoming buffer signal.

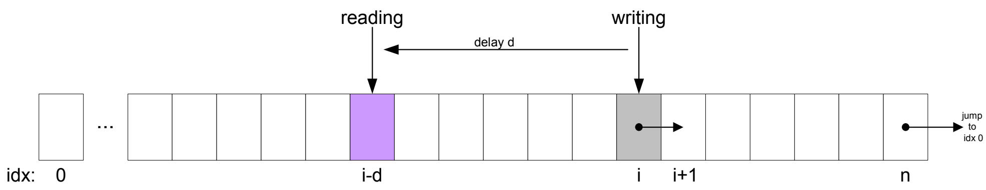


Figure 11: Ringbuffer

The reading point position in the circular array has to be calculated. This is normally done using a modulo operation for the pointer position in the array, which is sufficiently fast for many applications. In this case, however, the high amount of loudspeaker buffers with multiple pointers would cause a high CPU runtime. By implementing a *power of two* algorithm the total runtime can be diminished. The power of two algorithm works with the much faster operations *SHIFT* and *AND*, instead of *DIVISION* and *MODULO*. For a better understanding a simple example of the two mathematically equal algorithms is given. The power of two algorithm works only for divisors which can be rewritten as a power of two. Table 4 shows decimal numbers and their binary conversions. In this example a number $X = 247$ is divided by $D = 4$ which results in 61 plus a rest of $\frac{3}{4}$. In the binary range of numbers this result can be achieved similarly by *DIVISION* and *MODULO*. The second way is to shift X to the right by two bits (because 4 is the divisor) which results in 00111101. The rest of the equation is achieved by adding the divisor to the number X. [Hor12]

	number X	divisor D	result	modulo
decimal	247	4	61	3
binary	1111 0111	0000 0100	0011 1101	0000 0011

Table 4: Calculation example.

$$\begin{aligned} X/D &= 0011\ 1101 && \text{division} \\ X\%D &= 0000\ 0011 && \text{modulo} \\ X/D = X \gg 2 &= 0011\ 1101 && \text{shift} \\ X\%D = X\&0000\ 0100 &= 0000\ 0011 && \text{and} \end{aligned}$$

Implementing the delayed reading for the loudspeakers gives the possibility to alter the perception of the room size by introducing the parameter *delay factor* q_{dfac}. The delay factor is multiplied with the calculated delay. Positive values bigger than one result in a widening of the room. Taking a closer look at the delay determined by the system and the setup, it can be seen that the delay consists of 3 parts. On the acoustical path the delay is defined by the distance between sound source and microphone d_{sm} as well as the distance between microphone and loudspeaker d_{fix} [8]. Using the simple RMS algorithm (chapter 2.4.2) for localization, the position of the sound source is equal to the microphone position. Hence, the delay from sound source to microphone is not detectable, but it still takes part in the overall delay. Also, the latency of the system adds another delay d_{lat} on the electric part of the system. Compare with figure 12.

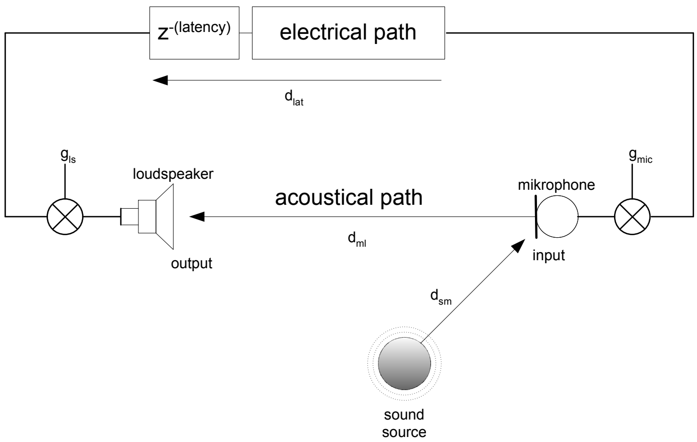


Figure 12: Delays.

As a result of the delay structure a bias of the minimum delay value is set by $d_{sm} + d_{lat}$. This means q_{dfac} only affects d_{fix} and a expansion of the room size can be achieved. Using values $0 < q_{dfac} < 1$ shortens the distances between microphone and loudspeaker virtually, but a reduction of the perceived room size is not possible. The room itself keeps it's acoustic qualities and defines the lower bound of perceived room size. Nevertheless, with a delay factor of $q_{dfac} = 0$ the feedback scenario described in 3.4.1 can be recalled, where no delay between microphone and loudspeaker is present $d_{ml} = 0$:

$$d = d_{sm} + d_{lat} + q_{dfac} \cdot d_{fix} \tag{30}$$

we also define

$$d_{ml} = q_{dfac} \cdot d_{fix} \tag{31}$$

As described in chapter 3.4.1 the arrangement of the setup forms a feedback loop system. By tuning the delay factor q_{dfac}, system attributes and feedback resonances can be altered and controlled. Introducing a delay factor $q_{dfac} > 1$ the resonances are at lower feedback frequencies. This is consistent with the calculation of the room modes in the model of a rectangular room, where the resonance frequency is indirect proportional to the room dimensions:

8. compare with chapter 3.4.3 (calculating the distances) and 3.5.2 . $_{fix}$ corresponds to the calculated distances from microphones to the fixed positions of the loudspeakers.

$$f_{n_x,n_y,n_z} = \frac{c}{2} \cdot \sqrt{\left(\frac{n_x}{l_x}\right)^2 + \left(\frac{n_y}{l_y}\right)^2 + \left(\frac{n_z}{l_z}\right)^2} \tag{32}$$

object	resonance frequency	order of mode	length	dimensions
symbol	f_{n_x,n_y,n_z}	n	l	x,y,z

Table 5: Symbols for calculating the resonance frequency.

3.4.3 Path scenario

As mentioned in chapter 3.2 the setup was tested with white noise and played back consecutively by all loudspeakers. This setup test was meant to control the function of the hardware and to detect any errors in single channels, like wrong patching. Still, the produced sound was found to be very interesting, because the white noise formed paths through the room, by jumping from one loudspeaker to the next one. The order of the loudspeakers played back followed the numbering of the outputs from 0 to 47, hence, always the same path was followed. Yet, due to the different orientations of the loudspeakers and the system, the perceived sound is a filtered version of the represented white noise in dependency of the listener's position. Therefore, the room, the position of the listener as well as the orientation of the loudspeakers shape the signal. Especially when standing in between of all loudspeakers in the installation space, the changing sound source positions are clearly perceived. Yet, on the one hand sound from loudspeakers nearby the listener can be easily detected, but on the other hand loudspeakers further away become indistinct in their location. Listening closely to the sound source and tracking it's position, one can hear that it follows a path through the room. The produced sound is completely different to the outcome of the feedback scenario as well as the delay scenario. In the feedback scenario one can explore a part of the frequency domain of the system by listening to the resonance tones. In the path scenario the variable is time. The system is presented from a different point of view . For us, these effects were a worthwhile experience and the decision was made to further develop the scenario and use it in the installation.

The concept of the installation is to work with sound that is produced in the room. Therefore the ideas generated by using white noise as an output signal were transferred to use sounds captured by the microphones. To avoid clicks when the representation of the signals change from one loudspeaker to the next, a signal amplitude fading (compare with 2.1.3) was used to prevent discontinuities.

Path definition: In order to produce different paths depending on the position of the sound source, a similar algorithm as used to assign the microphone signals to the loudspeakers was implemented (compare with chapter 3.3). The only difference here is that now the assignment is defined between loudspeaker to loudspeaker, or as we call it: loudspeaker to neighbours. Again, the idea is to connect one loudspeaker with it's two nearest neighbours and pass the signal. This means the distances for every loudspeaker to all loudspeakers is calculated. This results in an array of $48 * 48 = 2304$ distances, where the distance for loudspeaker i to itself is set from 0 to a high value, so as not to be selected by itself and cause interior feedback.

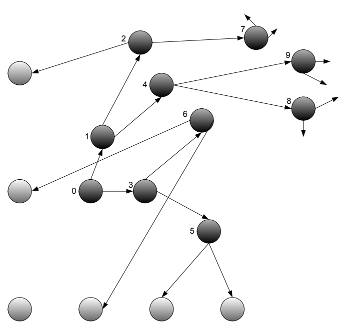
In figure 13 different solution for the assignment of neighbours to loudspeakers are shown. The illustration uses spheres to depict loudspeakers as masses, like in the physical modelling. The black spheres represent the loudspeakers, where the arrows show the assigned neighbours. The brighter spheres represent the helper masses (compare with chapter 2.3). Helper masses do not pass the signal to any other masses, hence, the signals arriving at a helper mass vanish. The four depicted loudspeaker positions and the corresponding algorithms will be described in the following part of this paper.

In the ideal case, one signal is passed from one loudspeaker to it's neighbours and distributes in the whole room, without repeating of one loudspeaker. Unfortunately, once a loudspeaker cannot be repeated, it is likely that big jumps from one loudspeaker to the next unassigned neighbours occur. Therefore, no clear paths are defined and the sound source seems to jump randomly in space. It is not possible to detect a path or a preferred moving direction. In figure 13a one can see that the distances of selected neighbours increase with higher iteration.

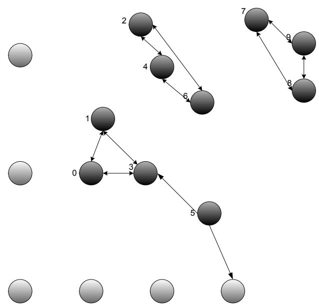
In order to avoid this problem another algorithm can be implemented, where a repetition of loudspeakers is allowed. Yet, due to the random distribution of the loudspeakers loops or "zones" are formed between closely located loudspeakers. Here the signal is circulating in the loop until it vanishes. Hence, the signal will enter the nearest loop to the sound source position and stay there. Of course, some loudspeakers work as connectors (compare with loudspeaker 5 in figure 13b) between certain loops and a spreading of the signal is not impossible, but this can be seen as an exceptional case. In other words, the signal distribution is not equiprobable, but some loudspeakers are less likely to be passed a signal than others. This effect can be seen in figure 13b.

Revised loudspeaker array model: Facing the aforementioned problems we started to rethink the positioning of the loudspeaker either in the real world as well as in the loudspeaker model for calculating the distances. The goal was to get a equiprobable distribution of sound from one loudspeaker to another, that forms a perceivable path through the installation space, as well as to define a regular sound field, without loops. The irregular loudspeaker positioning also creates problems for the other scenarios. For example the feedback loop system is more sensitive against resonance tones, which is ascribable due to the circulating of the signal in loops. Another kind of feedback occurs in the delay scenario, where no resonance tones are building up, but the overall signal level increases drastically. One way to work around these problems would be to find a set of parameters that would minimize these effects. We tried to find a more general solution for all these issues by revising the loudspeaker positioning.

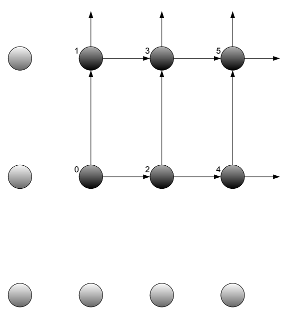
If the loudspeaker positions are regularized so as to become a symmetrical structure and a equiprobable path selection is possible, two approaches for the regularization come to mind. The first one is to change the loudspeaker positions and form a regular array in the installation. But this would not meet with the fundamental idea of the installation (compare with chapter 1.1 and 3.2.1). The second idea is to regularize the model of the loudspeaker positions and leave the real loudspeaker positions randomly distributed. In this case the selection criteria, hence, the distances between loudspeakers, are chosen to be equal. Therefore, the algorithm does not choose certain areas or loudspeakers more likely than others anymore. Nevertheless, the installation itself forms a random environment where the model is not an exact representation of the system. On the down side of this approach, the regular loudspeaker array model forms a regular path. This means, because of the equal distances between loudspeakers, neighbours are selected by a simple pattern, where the next loudspeaker in horizontal and vertical direction is chosen. A few exceptions occur at the border of the array. The path over the whole array follows two uniform directions (Figure 13c). For all algorithmic solutions, except the solution where repetition is allowed, the starting point affects the spreading of the signal the most.



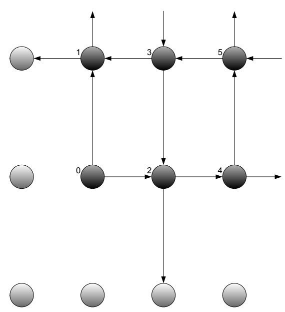
(a) Random loudspeaker position without repeating of neighbours.



(b) Random loudspeaker position with repeating of neighbours.



(c) Regular loudspeaker array with algorithmic assignment.



(d) Regular loudspeaker array with hardwired assignment.

Figure 13: Loudspeaker positions seen from above with different solution for the assignment of loudspeakers to their neighbours.

Finally the algorithmic approach was replaced by a hardwired solution, where a simple geometrical pattern is used to decide to which loudspeaker the signal is routed to. This maximizes the control of the distribution of the microphone signals regarding the number of loudspeakers and the positions. Also special cases like the routing of microphone signals from the environment to special loudspeakers in space are done in a more flexible and faster way. In fact, this process can be seen as a patching of signals from sources to sinks, like it is commonly done in sound engineering.

In this case the nearest neighbours are selected to prevent a dependency on the starting point and a previleged direction path. Also repeating is allowed, but no loops, where the signals circulate, are formed. Yet, all loudspeakers are connected to two neighbours and form a network where every discrete point of loudspeaker position can be reached with the same probability. Nevertheless, the starting point defines how long it will take to reach a specific position [9]. This means when starting for example at the center of the array, the signal will be passed from one loudspeaker to the next one, over the whole field. The pattern uses a *L*-structure, that is mirrored on the vertical axis for the next loudspeaker in vertical direction and mirrored horizontally for the next loudspeaker in horizontal direction. See also table 6. Figure 13d shows the final version of the loudspeaker assignment using a hardwired solution.

loudspeaker	1^{st} neighbour	2^{nd} neighbour
0	1	2
1	3	helper
2	3	4
3	1	2
4	6	5
5	...	...

Table 6: Hardwired loudspeaker assignment.

In figure 14 the assignment of loudspeakers to their neighbours for all 48 loudspeakers plus helpers is shown. Two exceptions to the former described *L*-structure pattern can be seen for loudspeaker 11 and 42.These loudspeakers send their signals to the corner loudspeakers instead of to the helpers where the signal would vanish. This is done to maintain the signal flow.

9. Also the amplification which is multiplied to the passed signal defines whether all points in the array can be reached or not. Where an amplification equal to one means no loss, smaller than one the signal strength decreases and amplifications bigger than one will lead to an increasing overall level.

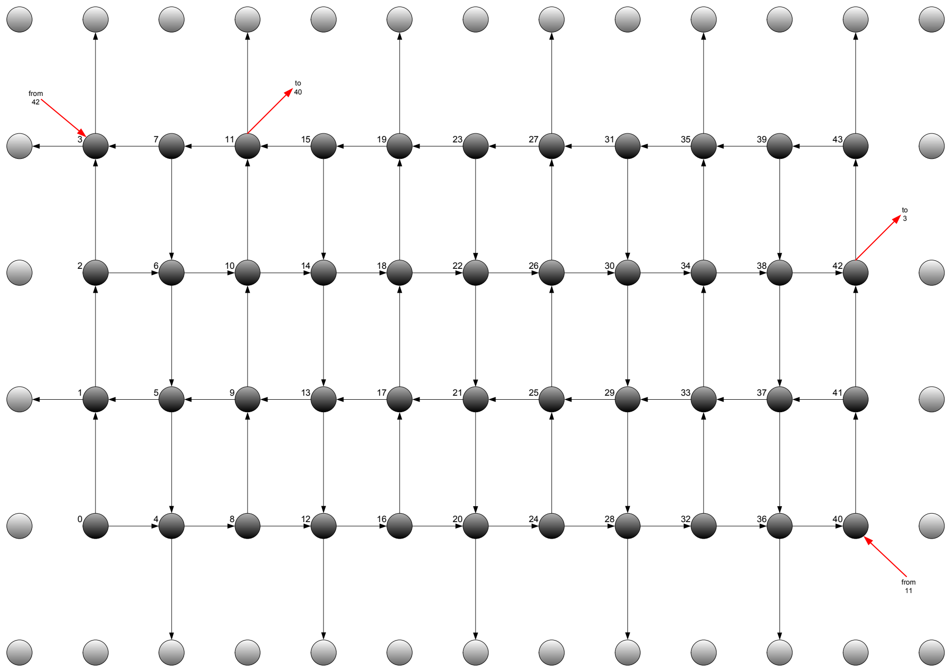


Figure 14: Regular loudspeaker positions with hardwired solution for the assignment of loudspeakers to their neighbours seen from above.

3.5 Merging the scenarios

The definition of the three scenarios (i) feedback, (ii) delay and (iii) path were the main system outcome explorations used as the starting point for further development. Although the scenarios do have certain areas of intersection, they are independent in many ways and their main control parameters are different. The question was how to construct a system that would be able to change and merge the scenarios in a suitable way. Solving this task can be done in multiple ways. One solution would be to change the system behaviour over time with a scheduled routine. Also a randomly merging of the scenarios would be possible, where depending on a randomly changing parameter decisions about the system state are made. A third possibility is to create zones in the installation. In this case the scenarios would be separated by invisible borders, defining the zones. Nevertheless, defining the areas as well as distinct and perceivable borders is a quite difficult task. Also, the organic behaviour of the installation gets lost due to separation.

The installation idea uses interaction as a main part of it's concept. Therefore, priority was given to an interactive solution, whereas time dependent or random distributed solutions were avoided. Also the approach using zones was tested, but because of the aforementioned problems it was not used for the final installation. The next section will describe the interactive solution, creating an organic installation scenario.

A physical modelling approach was chosen to merge the scenarios in a flexible application. The main advantage using physical modelling and *rattle* is given by it's capability of working in realtime. This allows altering of parameters while the system is running, without stopping and recompiling the whole algorithm. The implementation of the model will be explained in chapter 3.5.1.

3.5.1 Physical modelling of the system

In chapter 2.3 the two fundamental elements of physical modelling mass and link were defined. The equivalent to the loudspeakers and microphones in the real world are masses in the model. Every object (loudspeaker, microphone) corresponds to one mass in the model. The link defines the interaction between the masses. Therefore, on the one hand the link can be seen as the signal routing from one object to the next (microphone-loudspeaker, loudspeaker-loudspeaker) and on the other hand as the physical component defining the type of force (spring-like, gravitational) between the objects. There is a valid analogy for real world loudspeakers and microphones modelled as masses, which are concrete objects in both worlds, as well as for the intangible audio signals modelled by the links. This connection between the two worlds was chosen because of its plausibility and the graphicness of the simulation.

Thinking about the system's set up, two arrays of objects have to be modelled. The 20 microphones form a regular array of 20 masses at their corresponding positions[10]. The array is centered around the origin in the horizontal plane, but elevated to the corresponding height of the hanging microphones. The 44 loudspeakers follow the regular array shown in figure 14 and not the exact position of the distributed loudspeakers in the room (compare with chapter 3.4.3 *revised loudspeaker array model*). The array lies on the horizontal plane through the origin. The microphone array is fixed at it's position and no forces are exerted from other objects. The force of every microphone corresponds to its input signal. The loudspeaker masses move by these applied forces, but the initial positions are stored as fixed positions in order to be able to reset the loudspeaker array.

Let us discuss these matters in more detail. We defined, that the microphone array is at a fixed position and the microphones/masses are connected to their nearest two loudspeakers. Meaning, signals are passed from the microphone to the loudspeakers. The loudspeakers are connected as well, following the hardwired solution shown in figure 14 and they are not fixed at their positions. The connections between objects correspond to the linking element in physical modelling. This implies, only for these connections forces between the masses exist. Without any input at the microphones, all masses are stable at their equilibrium positions. This is achieved by exerting a small force to the loudspeaker masses, that keeps it at the stable position. In other words, the mass is surrounded by its field keeping it at the equilibrium position (Compare with chapter

10. In the installation space 16 microphones were installed as a regular array in room 1 and 4 microphones were placed in room 2. For the model a regular array of 20 microphones was used. Compare with chapter 3.2.2

2.3.5: field type 1). Also attrition [11] has a positive effect and prevents masses to react sensitively to small forces. The input of the microphones amplifies the forces of the mass and acts on the connected objects. This means a high input at a microphone effects the system the most. The positive force pulls the two connected loudspeakers towards the microphone. These two loudspeakers transmit the exerted forces to their two neighbours, via the links, which are modelled as spring-like forces. As a result the introduced forces propagate from one loudspeaker mass to the next and a wave is produced in the whole loudspeaker array. The wave propagation needs some time to develop, because of the inertia of the objects. Also attrition limits the perpetuation of forces. For steady acoustic signals at static points in the installation space, the loudspeakers would stabilize at a equilibrium position. In practice this case is not likely to appear, because of the time variation of sounds. Therefore, in this case the loudspeaker array is always moving. Nevertheless, using a set of parameters the wave propagation can be controlled or even turned off.

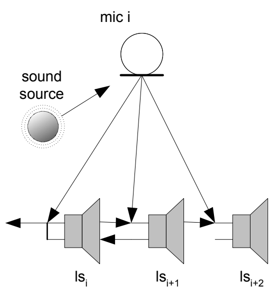


Figure 15: Minimal example of microphone to loudspeaker to loudspeaker signal routing.

In capter 2.3.3 the concept of physical modelling using fields was described. The parameters of the masses define the fields and are separately stored for every mass. We distinguish between a microphone field that acts on the link of microphone masses to loudspeaker masses and two fields that act only on loudspeaker masses. Loudspeaker field 1 keeps the loudspeakers at their position in the array, whereas loudspeaker field 2 defines the interaction between loudspeakers. Following parameters are used and have to be fine tuned to perform in the desired ways:

11. Equation (6) shows the dependency of attrition for forces.

		symbol	description
universal	loudspeaker	s-mass	mass of loudspeakers
		s-att	attrition for loudspeakers
	microphone	m-mass	mass of microphones
mic-ls	field type 2	m-k	spring rate for microphones
		m-pow	defines the force type for microphones
ls	field type 1	s-k	spring rate to keep ls at their positions
		s-pow	defines the force type to keep ls at their positions
	field type 2	ss-k	spring rate for forces from loudspeaker to loudspeaker
		ss-pow	defines the force type for loudspeaker to loudspeaker

Table 7: Symbols and description of the used parameters.

It can be seen that the forces (defined by the field attributes) are using basically the same set of parameters, but a separation was made to be able to use different values.

Using the highest input signal: 20 signal inputs introduce a huge amount of data and a very complex and complicated structure of forces will be applied to the physical model. In order to simplify the model, only the highest input signal measured by the RMS algorithm described in chapter 2.4.2 is taken into account. This is done by calculating the RMS values for all microphones, detecting the highest value and setting this index (microphone number) to 1, whereas all other indices are set to 0. This leads to a vector (*float RMS_max*) of 20 elements with all zeros except one. As noticed before, the goal is to use only the force introduced by the detected highest value. This is done by simply changing the spring constant of the microphone number. By increasing k for the microphones the force increases as well. The vector *RMS_max* is multiplied with a constant parameter, that defines the amount of the physical modelling impact, and is subtracted of the stored spring constant of the mass. To make it more clear, here is the equation and a simple example, where microphone 1 has the highest input:

$$k = m\_k - RMS\_max \cdot phy\_mic\_k \tag{33}$$

$$\begin{bmatrix} k_0 \\ k_1 \\ \vdots \\ k_{20} \end{bmatrix} = \begin{bmatrix} m_k \\ m_k \\ \vdots \\ m_k \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \cdot \begin{bmatrix} phy\_mic\_k \\ phy\_mic\_k \\ \vdots \\ phy\_mic\_k \end{bmatrix} \quad \text{where } m_k \text{ is negative, default } -0.01 \tag{34}$$

$$F = -k \cdot x \tag{35}$$

The multiplication with the parameter phy_mic_k allows to change the intensity of the alteration caused by the physical model and also to switch it completely off. GELESN BIS HIER

3.5.2 Application of the physical model

So far a model of the microphone and loudspeaker set up has been defined. The implementation as a physical model allows to run a simulation, where forces between objects are calculated corresponding to the input signals. The signals are used to interact with the model. We will now describe, the application of the model and the connection to the former introduced scenarios.

In chapter 3.3 and chapter 3.4.3 the calculation of distances between microphones and loudspeakers, as well as loudspeakers and loudspeakers is described and used to find a routing for the input and output signals. Running the physical model, the position of the loudspeakers changes, which leads to new distances between the connected objects. Therefore, also a new set of delays can be calculated for every loudspeaker. These delays are used to define the reading position in the ringbuffer of every loudspeaker (compare with figure 17). As a consequence, the delays are changing over time according to the simulation, whose motion depends on the input signals. Yet, the new delays do not affect the signal routing. This means that the signal routing and the chosen connections are still fixed.

Speaker delay parameters: We wanted to maximize the flexibility as well as the adjusting possibilities for the delays. Therefore, more parameters were necessary, which lead to a slightly more complex calculation equation for the delays. Taking a closer look, one can see that the following structure is basically the same as for the spring constant in equations (33) to (35).

$$d_{ml} = d_{fix} \cdot q_{dfac} + d_{calc} \cdot q_{dphy} \tag{36}$$

Equation (36) calculates the delay for the reading position in the ringbuffer[12]. In the first term the fixed position of the loudspeakers d_{fix} is multiplied with a factor q_{dfac} to virtually widen or reduce the overall room size. The second term consists of the new calculated delays d_{calc} weighted by a parameter for the effect of the physical modelling q_{dphy}. It should be mentioned that d_{calc} is the time, sound would need to cover the distance from the fixed position of the loudspeaker to the currently modelled position. A graphical overview is given in figure 16.

Merged scenarios: The three main scenarios can be separated by the delay parameter:
- feedback scenario: no delay between microphone and loudspeaker, which leads to a overall delay of $d = d_{sm} + d_{lat} + d_{ml}$ with $d_{ml} = q_{dfac} \cdot d_{fix} = 0$ (See chapter 3.4.2) The delay has a value in the range of a few samples.
- delay scenario: distances between source and sink are taken into account, widening or reduction of virtual room size with multiplicative factor. The delay has values in a

12. Equation (36) is an modification of $d_{ml} = q_{dfac} \cdot d_{fix}$ defined in chapter 3.4.2

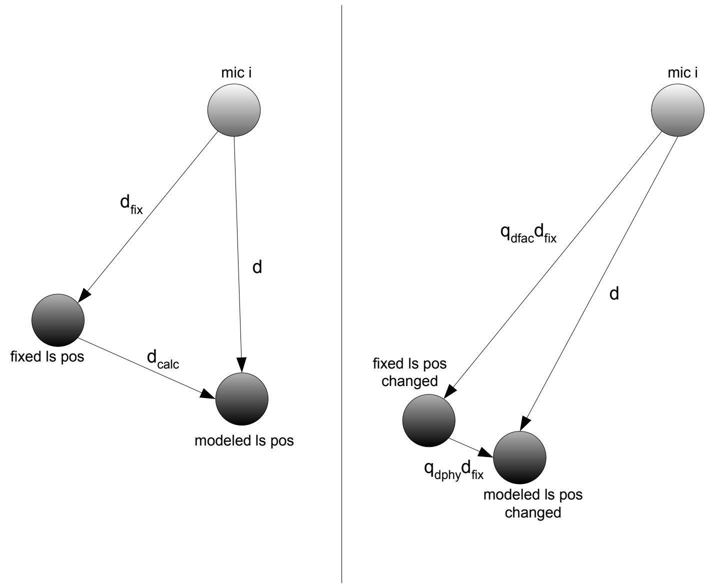


Figure 16: Calculating the delay d with and without alteration by the factors q_{dfac}, q_{dphy}.

range of 50ms to 200ms, which form echoes.

- path scenario: signal is passed from one loudspeaker to two others and played back with delays corresponding to the distances.

By introducing the distinction of the scenarios by the amount of delay, it can be seen, that the physical model switches or merges the scenarios depending on the input signals and on its state.The system state is the current behaviour and movement of the system. Hence, speaking in terms of RMS values, we can define the following cases corresponding to the scenarios:

- **no or very low RMS input signals (feedback):** Low input signals cause little forces applied to the physical model, which leads to no or little movements of the loudspeaker masses. They are kept at their fixed positions. Still, the delays of the distances between source and sink (introduced by the delay scenario) are present, which were not for the feedback scenario. Being in the room, one can hear that the perceived room size is widened. In other words the room appears to be more reverberant. Yet, feedback tones still emerge, only a longer period of time is needed until the closed loop amplifies the fed back signal. At the point where feedback tones gain a certain amplitude level (determined by the field that keeps the loudspeakers at their initial position) the applied forces between microphone and loudspeakers are big enough to move the loudspeakers. This causes an alteration of the calculated delays, and therefore the position of the reading point in the buffer. Hence, the system attributes as well as the resonances are changing and the feedback vanishes. In this case the feedback scenario kills itself by its own growth.
- **middle RMS input signals (echo):** If signals with middle RMS values occur in

the room, the forces exerted to the objects change the modelled loudspeaker position in the physical model simulation. Also, the signals is spreading in the room because it is passed from loudspeaker to loudspeaker. As described earlier the loudspeaker positions are changed which results in an alteration of the reading point position in the loudspeaker buffer. The signal is played back as echoes from different loudspeaker positions.

- **high RMS input signals (path):** Signals with high RMS values, like clapping or whistling, will force a vast alteration of the loudspeaker mass positions. At first the high force pulls the loudspeakers towards the microphone. Next, the affected loudspeakers pull their two linked neighbours towards their positions. These two steps lead to big distances between the equilibrium and the current position for every loudspeaker. Hence, the result is a high delay value for the reading point position in the loudspeaker buffers. Therefore, the signals played back by different loudspeakers can be distinguished and paths become noticeable.

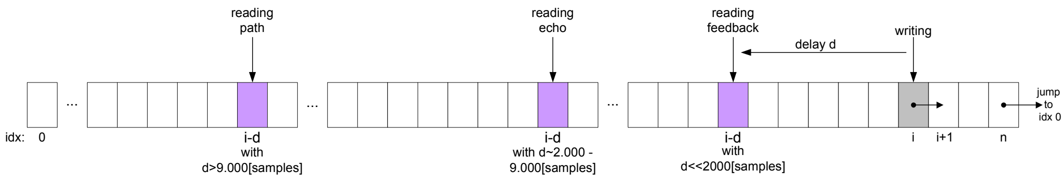


Figure 17: Reading point positions in the buffer corresponding to the used delay ranges for the three scenarios.

The echo and the paths scenario are related closely, but they do not interfere with each other. On the one hand this is because of the different range of delays (see figure 17)used in the scenarios and on the other hand the amount of their appearance can be controlled with two separate amplification parameters. In simple words, the amplification parameters are multiplied to the signals read from the microphone and stored to the loudspeaker buffers. They determine the signal level for the output. Nevertheless, the overall signal level influences the feedback scenario as well.

3.5.3 Pitch shifting

As seen before, the reading positions in the buffers jumps with the alteration of the calculated delays. Playing back a signal with a changing buffer reading position poses some problems. With every jump in the ring buffer a *click* will be heard. Crossfading between the previously and the currently read signal avoids the creation of clicks. This is implemented by using an integrator, in our case the *lag*-function described in chapter 2.1.3.

Another problem is the effect of pitch shifting the signal. Pitch shifting occurs because due to the changing reading positions. By jumping from one position to another, some samples are not played back. This results in either in a higher playback speed and therefore in a higher pitch or in a replaying of an earlier signal part. This means pitch shifting is only possible upwards.

Another way to explain the pitch shifting uses the *Doppler effect.* As explained, the delay corresponds to a distance. By changing the delay, the signal is virtually moving towards to or away from the listener's position. For a constant position the distance does not influence the pitch of a signal, but a moving sound source results in the *Doppler effect*. The Doppler effect occurs, because the sound wave is traveling with speed c[13] in all directions, but the moving source reduces or widens the distances between the peaks of the waveform. Following, the wavelength between two peaks with similar signs is changed, which is indirect proportional to the frequency. Moving towards the listener's

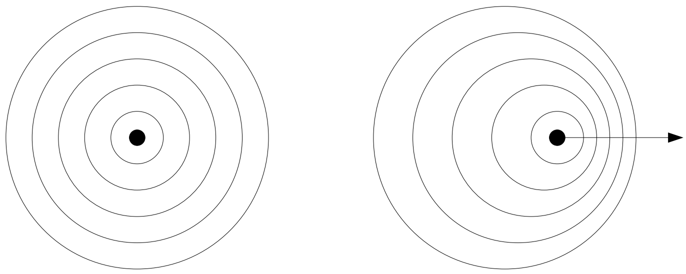


Figure 18: The Doppler effect.

position or reducing the delay results in higher pitches of the signal; moving away or increasing the delay results in lower pitches. Steadily changing the delay for the reading position introduces a virtual movement and therefore velocity to the sound source, which is equal to the Doppler effect [Moo90].

3.6 Loudspeaker output calibration

The transfer function of a system is a way to visualize the system's attributes. In the case of a feedback loop system the transfer function has clearly pronounced resonances. These resonances cause feedbacks at specific frequencies. As described in chapter 3.4.1 the ideal system state to produce feedback is the unstable balance. Still, the resonances are present and easily result in feedbacks with static frequencies. One reason is the predominance of certain loudspeaker and microphone pairs, which are more likely to produce feedbacks than others. Especially for the irregular positions of the loudspeakers, where some distances between loudspeaker and microphones are smaller. Variable feedbacks depending on interaction were desired for the installation. For this reason two ways to diminish the predominant resonances were implemented.

13. see eq.(27)

3.6.1 Loudspeaker output amplitude calibration

In order to reduce the effect of different distances between microphones and loudspeakers, the simplest way is to adjust the output amplification for every loudspeaker. Hence, amplitudes of loudspeakers close to a microphone are reduced and amplitudes of loudspeakers further away from microphones are amplified. In this way, other resonances are shifted upwards to the threshold of unstable balance and the set of occurring feedback frequencies is widened. In the case of one microphone and n speakers the method is straight forward, but with m microphones and n loudspeakers it is a quite complex task.

The loudspeaker output amplitude calibration has to be made once after setting up the installation. This is done by playing back *white noise* for 2 [seconds][14] for every loudspeaker separately with a break in between of 2.5 [seconds]. During the playback the algorithm described in chapter 2.4.2 calculates the RMS value for every microphone m. In this application the *blocksize* has been changed to $4096 \cdot 16$ [samples] and the *hopsize* to 1024 [samples]. The processed measurement data is stored to a file for further reference.

The next step is to find the maximum RMS value captured by one microphone for every speaker n (RMS_n) as well as the absolute maximum RMS value (RMS_{max}) of all loudspeakers and microphones. Table 8 gives an overview of the data. The highest value of line 1 is RMS_1, line 2 is RMS_2 and so on. The highest of these values is called RMS_{max}.

		microphone							
		1	2	...	20				
ls	1	$RMS_{1,1}$	$RMS_{1,2}$	...	$RMS_{1,20}$	$\Rightarrow$	RMS_1		
	2	$RMS_{2,1}$	$RMS_{2,2}$		$RMS_{2,20}$	$\Rightarrow$	RMS_2	$\Rightarrow$	RMS_{max}
	$\vdots$			$\ddots$	$\vdots$		$\vdots$		
	44	$RMS_{44,1}$	$RMS_{44,2}$	...	$RMS_{44,20}$	$\Rightarrow$	RMS_{44}		

Table 8: Overview of the RMS values.

The amplitude correction can be calculated for every loudspeaker by

$$Acorr_n = \frac{RMS_{max}}{RMS_n} \tag{37}$$

and stored in a separate file, that is loaded into the installation code. The corrections are directly multiplied at the output of the *rattle* environment. Taking a closer look at the calculation of the corrections, one can easily see, that all outputs are amplified. The only exception is the output of the loudspeaker producing the RMS_{max}. Hence, the overall *master volume* has to be reduced in order to reach the state unstable balance.

14. In the code the time has to be given in microseconds $2 \cdot 10^6$ [μs]

3.6.2 Loudspeaker output frequency calibration

The second approach to reduce the major resonances, is to cut the specific frequencies using an equalizer. This process is commonly used in sound engineering, yet, here a full parametric peak filter is used instead of a 31-band graphical equalizer. The algorithm behind the 2^{nd} order peak filter is described in chapter 2.1.1. The filter is implemented in all outputs. Nevertheless, it was sufficient to cut only the major resonance for all outputs. This means only $2683Hz$ were cut by $-11dB$ and a Quality of $Q = 2$. This second approach was done after the loudspeaker output amplitude calibration.

3.7 The audio callback - output signal routing, reading, amplification, modulation

In this chapter the final audio callback used in *rattle* will be described. The audio callback defines the signal which is passed to the DA converters and consecutively to the loudspeakers. This function contains the "dry" microphone signals, signals sent from neighbouring loudspeakers as well as the alterations by the physical modelling. It can be seen as the core function of the installation, bringing together all functions, solving approaches and decisions concerning the final installation implementation. Previously, these separate parts of the algorithms were described and now we will continue with consolidation.

As alluded above, three major signal branches are weighted and summed up to form the output signal. These signals are the

(a) direct microphone signal: read without delay

(b) self microphone signal altered by physical modelling: read with delay corresponding to distance between microphone and connected loudspeaker

(c) signal sent from neigbouring loudspeakers altered by physical modelling: *other*

Equalization, limiting, loudspeaker output amplitude correction (chapter 3.6.1) and sending signals to the next neighbour is performed iteratively for every loudspeaker in the audio callback *loop*-function [15]. Figure 19 shows the block diagram of the implemented code and table 9 gives an overview of the used signals as well as the used parameters, which can be categorised into gain factors and delays.

15. The *loop*-function is an implementation of a simple *for* statement, with lower computational costs, which is essential for computing audio applications.

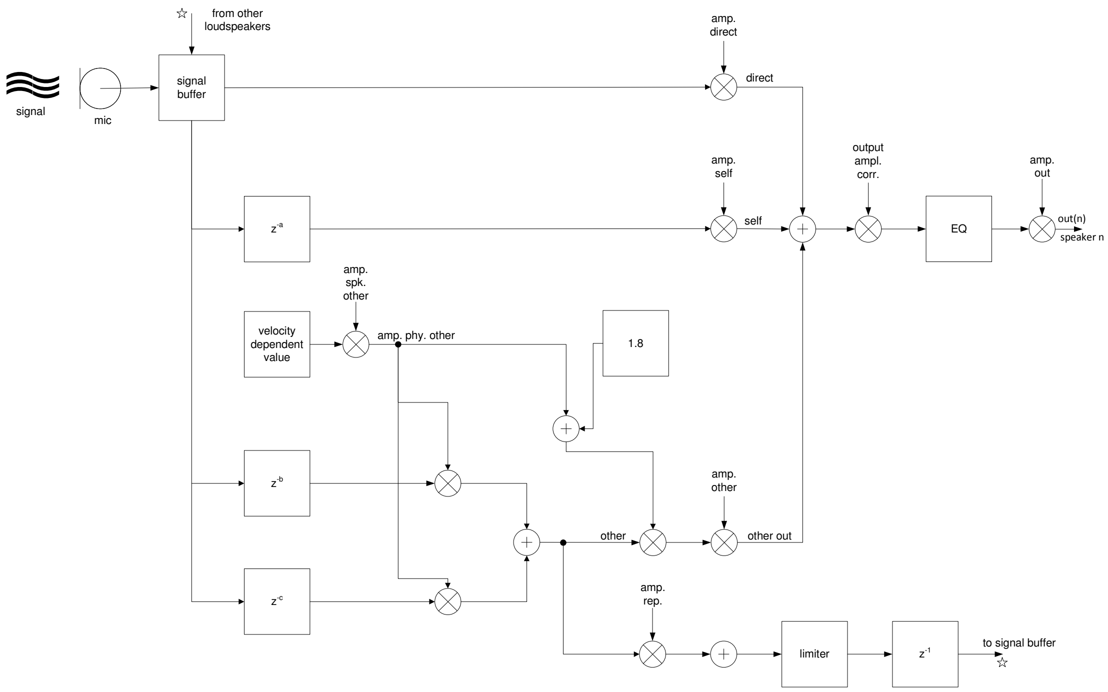


Figure 19: Blockdiagramm of the audio callback.

	parameters			
	name	**code**	**branch**	**description**
gain factors	amp. direct	amp_direct	direct	weights the amount of direct mic signal
	amp. self	amp_self	self	weights the amount of delayed mic signal
	amp. spk. other	phy_spk_other_amp	physical modelling	weights the amount of physical modelling impact
	amp. other	amp_other	physical modelling	weights the amount of signal altered by the physical modelling mixed to the output
	amp. rep	amp_rep	physical modelling	weights the amount of signals sent to the signal buffer,hence to the next neighbour
delays	a	mik_del	self	defines the reading position in dependency of the physical modelling
	b	mik_spk0	physical modelling	defines the reading position for the signal sent from neighbour 1
	c	mik_spk1	physical modelling	defines the reading position for the signal sent from neighbour 2

Table 9: Overview of the parameters used to mix the output in the audio callback.

The delays a, b, c shown in table 9 are calculated in the following way using fixed positions as well as the changing positions introduced by the physical modelling. Out of these positions the time delay for a propagating sound wave can be calculated.

$$a = \begin{cases} d_{fix} + d_{stable} + d_{calc} & \text{if a>0} \\ 0 & \text{if a<0} \end{cases} \tag{38}$$

$$b = \begin{cases} d_{neighbour1\_fix} \cdot q_{dfac\_spk} + d_{calc} \cdot q_{dphy} & \text{if a>0} \\ 0 & \text{if a<0} \end{cases} \tag{39}$$

$$c = \begin{cases} d_{neighbour2\_fix} \cdot q_{dfac\_spk} + d_{calc} \cdot q_{dphy} & \text{if a>0} \\ 0 & \text{if a<0} \end{cases} \tag{40}$$

delay	**code**	**distance**	**see also**
d_{fix}	d_fix	microphone to fixed loudspeaker	eq. (31)
d_{stable}	d_stable	?=d_{fix}?	eq. (36)
d_{calc}	d_fix	microphone to moving loudspeaker	eq. (36)
$d_{neighbour1\_fix}$	$d\_$	loudspeaker to neighbour1	ch. 3.4.3
$d_{neighbour2\_fix}$	$d\_$	loudspeaker to neighbour 2	ch. 3.4.3

Table 10: Delays

In equations (38) - (40) the distinction for the values of a, b, c is made in order to guarantee a (theoretical) causal system, where only past signals can be read out of the signal buffer. Nevertheless, if values < 0 are used the reading position in the buffer is ahead of the writing position, which will not result in an acausal system for the implementation, because in this case only data stored at a previous time is used. The result is in a very long delay, but also effects like reverse audio playback may occur.

4 Perceived installation outcome

In the previous chapters we described the concept, the designing, the implementation and the development process of generating *Interstices - Zwischenräume*. Here, a short summery of the resulting installation is given, focusing on the perceptional aspects.

Once the installation setup in Forum Stadtpark was fixed, fine-tuning the parameters was an important and time consuming part of the work. As said before, we wanted to create every nuance of system outcome from subtle to flamboyant, depending on the interaction taking place in the room. The three major scenarios (i) feedback, (ii) delay and (iii) path should be clearly perceptible and distinguishable, yet, forming a homogeneous body of sound. Merging the scenarios was done with the help of the physical modelling where the interaction and therefore level and position of sound production was taken into account.

For us, the final installation provided several changing acoustic responses evoking a new auditory environment. Especially the widening of the room resulting from a longer reverberation time introduced by the delay scenario was well achieved. By entering the installation space little accidental noises already formed an impression of a bigger room size. We hoped that these small indications would stir up curiosity of the visitors to start to stomp their feet, to clap their hands, or similar action. Increasing the sound level of produced noises lead to louder and much more pronounced alterations of the room responses. Signals are passed from the nearest loudspeaker of the sound source position to its neighbours. A seemingly random path is formed spreading in the room forming a circular movement. Because of the variable reading speed in the buffer, the signal is pitched. The responses allow a wide interpretation and connections to noises made in natural environments can be found. The visitors fantasy directs his/her perception as well as the interaction.

Composing the feedback scenario and fitting it into the whole installation was found to be a quite difficult task. Only on rare occasions, with very low input levels, feedback frequencies could build up and eventually destroy themselves again. Also triggering feedbacks, by covering loudspeaker membrans needed a long time. Therefor and because of the interactive approach with immediate responses for all other actions, this part of the installation was pushed to the background.

The four loudspeakers in the paper heap played back sound captured at that moment outside of the building. By entering the paper heap these signals were shifted to the foreground of the attention. Scenes of the park were clearly recognizable (speech, playing children, water fountain, barking dogs,...). Nevertheless, one was forced to sit in the cave to hear these events and being not able to look out of the window and see the sound sources. Therefor, only the auditory source but not the visual source was detectable. The presented sound gave the impression of a reproduced soundscape. Also a second effect emerged, where one reckons that a window is opened. This open window effect occurs, when one stands in the installation space, looking into the park and hearing the matching sounds from inside the paper heap. The visual and the auditory source were detectable, but separated. Although the direction of the sound was different, it fit the perception. The sound was sensed as a reflexion in the room.

This small summery gives information about the understanding of the installation by the designers. In the next chapter it will be evaluated, how a small representative group of people, who were not involved in the development process, perceived and described the installation.

5 Evaluation

In this section we will describe the evaluation process, the applied method and the findings. The goal was to collect qualitative data about visitors perception and the developed *visitor*-system-model of the sound installation. From this data we tried to find a pattern or a regularity about the acquisition of knowledge about the system by visitors. All findings were amalgamated to eventually form a theory about perception and understanding of the installation and compared to the understaning of the installation by the designers.

5.1 Evaluation concept

In order to collect data of visitors perception, seven probands were invited to visit the sound installation and spend as much time as they liked in the installation space. Afterwards they were interviewed about their gathered experiences. The interviews were held as informal as possible, with the goal to give the proband the possibility to talk freely about his/her perceptions. Six questions were defined as a framework for the interview. The probands did not get any information about the installation before the interview[16]. The questions were formulated in a very general manner. The interviews were recorded using a simple audio recorder. Using such a device benefits the flow of conversation and it can be listened to and worked on at a later time.

5.1.1 Persons

In this paragraph we want to give further information about the probands. The following table covers the basic informations as well as the educational background[17] of the persons.

Name	age	gender	musical education	occupation
visitor 1	25-30	m	none	journalist
visitor 2	25-30	m	git(hs), sound engineering(u)	student
visitor 3	20-25	f	cello(c)	office admin.
visitor 4	20-25	f	violin, piano, vocals (c)	student
visitor 5	25-30	f	piano(u), sound engineering(u)	student
visitor 6	25-30	m	piano(c), sound engineering(u)	student
visitor 7	25-30	m	clarinet(hs), sound engineering(u)	engineer

Table 12: Persons

16. At least no information was provided by the interviewer, nevertheless, some probands read fliers or visited the homepage of *Forum Stadtpark* to gather information in advance.
17. nouveau of musical education given in brackets (High school (hs), conservatoire(c), university(u))

Questions: The questions aimed mainly on the subjects of auditive perception, interaction, conceptual understanding of the installation and personal cognition. As mentioned above the interviews were held in a quite informal way and tried to get as much information as possible without pushing the proband into specific directions. Nevertheless, help was offered once a proband got stuck in his explanations or when the conversation went off topic. Also, some more detailed questions were asked after the probands described their experiences. These questions were either to better understand the meaning of their answers or to point out a topic, the proband did not mention. Following questions build the framework:

1. What did you perceive when entering the sound installation at first?
2. Did you produce any sound on purpose and what was the reaction?
3. Did you get any kind of feedback when producing sound?
4. Have you been into the cave and what did you perceive?
5. Did the sound installation change over time?
6. Which attributes would you assign to the installation/the perceived room?

Videos: In addition, visitors $1, 2, 4, 7$ were filmed while staying in the installation space.

5.2 Interview evaluation

In this section we want to give an overview on how the 7 visitors perceived the installation, which expectations they had, which impressions emerged and how they tried to explain the system responses to form a conceptual model. In chapter 5.3 we will discuss these matters in a more general way using the grounded theory approach.

Chapter 1.2 gives an overview on the installation and pictures the visual aspects. As designers with a background on sound engineering, computer music, acoustics and psychoacoustic our main focus lay on developing the auditory aspect of the system. To cite the American sculptor *Horatio Greenough*[18] also in our case *form follows function*. The mounting of microphones and the connection via cables were carefully designed, but the function lead to the final decisions. With our main focus on sound, the visual aspect was shifted to the background. Nevertheless, the visual appearance of the installation can not be separated or neglected. Five of seven individuals answered corresponding to the visual appearance when asked what was the first thing they noticed when entering the room. Especially the paper heap, the loudspeakers and the cables hanging from the ceiling attracted the most attention. Two visitors mentioned the microphones. Only the answers of two visitors corresponded to the altered sound perception as the first impression of the room. After their first explanations all visitors were asked, if the auditory perception of the room was identical to its visual appearance. For all individuals the room appeared to be much more reverberant and bigger than it was expected corresponding to the visual appearance. One visitor explained, that the room itself was

18. later also used by various architects including Frank Lloyed Wright

perceived without a lot of reverberation ("dry") but the overall sensation of the system formed the impression of a big hall, but altered.

The interviews showed that the visitors had different approaches and expectations to the installation. These approaches were lead by the question of finding out *how it works?* and *how to start?*. Also a soundscape or at least some auditory presentation immediately when entering the space was expected. It can be said, that some time was required to explore the interactive aspect of the installation. In many cases, only after an incident that produces sound the visitors started to play with the responses given by the room. Therefore an *interaction threshold* can be defined. The *interaction threshold* separates on the one hand a none interactive behaviour of an installation visitor and on the other hand an interactive behaviour. Naturally, it depends on the proband how high the *interaction threshold* is set. It is also very likely that a visitor, who enters the installation alone and does not get any information, leaves without interacting or producing sound at all. This effect occurred several times, with visitors who were not involved in the evaluation.

Once it was understood that *the room immediately answers to your actions* visitors started to explore the installation and used the offered and also the hidden affordances. By producing sound they tried to change the responses and to develop a conceptual understanding of the system. Especially visitors with a background in sound engineering tried to explain and understand the process. Appearing echoes or delays were sensed by all probands. The descriptions used terms that can be grouped to attributes of time and direction. The time attributes included variation in time, enhancement/shortening or brocken playback. The directional aspect was described as propagation in the room, circular movement of sound, signals from all directions and different positions. Explanations about these incidents covered theories about dependencies of the position of sound source as well as the position of loudspeakers, dependency on action and different applied effects. Also a dependency on the frequency range of the signal was mentioned by one visitor. The different directions of the played back sound was clearly detectable for the probands, but the construction of paths was not clearly detectable. Especially, it was not possible to formulate a regularity which describes the path. All in all it can be said, that the overall impression of the systems behaviour was *not obvious* but *chaotic* or *confusing* at first glance and later a *habituation* to the behaviour and outcomes emerged.

Three visitors verbalised the pitch shifting effect. One visitor tried to produce a triad by consecutively playing the single notes on the whistle. The result was different than expected, but according to the explanation a deeper understanding was achieved, although it was not explicable by the proband.

A special case in the process of exploring the installation is given by the action of entering the paper heap. At some point[19] during the visit all probands crawled into the heap. Before, the visitors learned that only an action triggers a response of the room and by crawling into the paper heap, this rule is not valid any more. Finally, a

19. It has to be mentioned that it took some time until probands went into the paper heap. None of them immediately entered it. The entrance lay on the backside of the heap pointing away from the doorway of the building which maybe biased the action.

soundscape was played back and it could not be influenced by the probands. This fact lead for four visitors to not fulfilled expectations, which resulted in some kind of slight disappointment. The soundscape was recognised as a scene in a public location, even one proband described it as a scene in the park. Yet, only two visitors explained that the sound was captured somewhere in the park at realtime. Three visitors thought it was a previously recorded and reproduced soundscape. For the remaining two visitors the conceiving of their opinion was biased, by reading the information on the flyer of the installation and by hints by other people visiting the installation. For two probands the soundscape was inviting as well as relaxing. Sounds from outside the paper heap where not audible or hard to distinguish for five probands, mainly because they focused on the happening inside the cave. Nevertheless, for one proband it was clearly understandable but dampened.

As mentioned before especially sound engineers tried to find an explanation about the system's behaviour and once they found one satisfying solution, the exploration was soon stopped. Other probands did not focus on a technical approach and on defining a theory about the system outcomes. This group was interested in the emotional interpretation of the system outcome. Some probands described the system outcome in a more general way and defined the outcome to none sensory attributes. For example one visitor explained that silent or calm behaviour resulted in an inviting atmosphere, whereas loud signals resulted in aggressive and scary noises. It is also worth mentioning, that two visitors explained that they varied the level of produced sound, but the direct dependency of level and system bahaviour was only found for one proband (emotional interpretation: inviting/aggressiv).

Interesting when does habituation emerge? Why stop to explore?

5.3 Grounded theory

As alluded above, the goal of the evaluation was to develop a theory of how visitors perceive the installation and how the process of *getting to know and understanding the behaviour* of the overall system works. We applied the method of *grounded theory*, which uses qualitative data to perform theory research.

5.3.1 Basic information about grounded theory

Grounded theory can be seen as a strategy to perform theory research. It works with qualitative instead of quantitative data and theories are derived using the data and then illustrating characteristic examples ([GS67]; p.5). The data can be taken from several sources, from field studies and interviews, articles, newspapers, letters, video recordings or any other kind of qualitative documentary material. *...anything that might shed light on the area of questions under study.* [CS90] With grounded theory, the researcher has to begin a systematic search for important categories ([GS67];p.169), but he/she is also encouraged to flexibly use the data and generate a theory with an open mind for all emerging and relevant aspects. Yet, some procedures and canons as described in [GS67] and summarized in [CS90] have to be taken seriously in order to evolve a theory based on grounded theory research. These 12 canons also outline the process of generating theory and they can be used as a guide towards a successful research outcome. An example of applied grounded theory research is given by the paper on *Conceptualization of violin quality by experienced performers* by C. Satis et al. It shows the hierarchy of themes, concepts, subconcepts and attributes, as well as a schematic representation of description of a violin quality evaluation. We tried to follow this example and develop a similar representation and hierarchy for our field of study.

5.3.2 Analysis

For the analysis of the data the *open coding* approach was used. In this process, the data is iteratively perused and incidents, events, happenings, (inter-)actions, etc. are marked, labeled and conceptualised. Differences and similarities are examined and concepts, categories and subcategories are formed. Yet, also *selective coding* was used to aim the research into the desired direction of examining action/interaction/perception of visitors.

5.3.3 Themes

Analysing the data, following four major categories or THEMES emerged.

- EXPECTATION - goal
- ACTION
- FEEDBACK - sensory experience
- INTERPRETATION - emotional & cognitive

These THEMES can be seen as the states a visitors runs through for several times, while staying in the installation space. The major categories help to construct a theory of how the installation is perceived and how people act during their visit. Figure 20 shows an overview of the major concepts as well as their nested sub-concepts and properties. Also, table 13 provides more information about the classification, with respect to the hierarchic structure of the concepts.

THEME	concept	property	description, classification	#	# sum
INTERPRETATION	emotional		inviting(1), well-being(1,3), aggressive(3), scary(3), confusing(4),relaxing(4)	3	5
	cognitive	unclear	not obvious(1,3),habituation(1,4), chaotic(4),	3	
		clear	clear(2,6),	2	
	explanation		depending on angle(1), regularity(4,6), alteration possible(5), effects(5), depending on action(5,7), room doesn't change(6,7), position(7)	5	
	conclusion		reverb-roomsize does not fit (4), perception changes(6)	2	
	HYPOTHESES		more people: hard to detect response(1,6),	2	
			other sound: not inviting impression(1),	1	
			tone quality (7), frequency(7)	1	
EXPECTATION	approach	predisposition	enttäuscht(1), how to start?(1), trying out(3), no expectations(4), how does it work?(5), feedback system(6), as expected(7), soundscape(7)	6	6
ACTION	accidential		drop jacket(1), steps(3,6), listen to single loudspeaker(7),	4	7
	inquistive		varying loudness? level(3,4), pitch(4), noises(4), scream/sing(5,6)	4	
	planned		triad(4), picking up paper(1)		
	initiated		rubberduck(1), objects(1,2,7),	3	
FEEDBACK sensory experience	visual	objects	loudspeaker(2,5), microphones(4,7), paper heap(3,4,5), cables(3,4,6,7), empty(3),	6	7
		roomsize		1	
	auditory	spatial	room: reverb(1,2,3,4), bigger(1,2,6,7), hall(1,2,5,6),	7	
			direction: propagating in the room(1,6), all directions(1,2,3,5), circular(2,4,6), position(4,5)	6	
		loudness	gain(1),	1	
		timbre	complex(4), not definable(4), distorted(5,7), neutral/subtle(6), intensive?(2,5,7), inerpreted sounds*	7	
		time	brockenly (1), moment(4), enhanced/shortened(6), varies in time(6), same sound again(1,3,4,6), delay(2,3,5,6,7),	3	
		pitch	pitch(1,5,6), transposed(5),	3	
	compatibility		roomsize,	3	

Table 13: Classification of concepts, categories and themes defined with the data of the interviews.

* these include:
· real: barking(1), noise related to water(1,2,4) speech(2,3,4), park(2,3)
· interpreted: humming(3), whistle(3)
· figurative: Native American tribe(3), ghost(3), animals(3,5)

Figure 20: Graphical representation of the classification of concepts, categories and themes defined with the data of the interviews.

The graphical representation of the classification of concepts, categories and themes help to build up a basic theory about the perception of the installation. Starting at the bottom of the graph, visitors expect some occurrence in the forefront. These expectations can be verbalised and clear defined or just a vague idea. This depends on the predisposition of the visitor. Entering the installation space is the first action performed. Immediately, sensory experiences[20] are made. The room is scanned visually, but also auditory feedback

20. here only the visual and the auditory experiences are examined. Yet, all human senses work closely together to collect data of the surrounding environment.

is collected and both sensory experiences are compared in order to construct and verify a picture of the surrounding environment. The collected data is interpreted and matched against the aforementioned expectation. This leads to fulfilled or unfulfilled expectations. The stage of interpretation can be separated into an emotional and a cognitive part with different percentage of allocation for every individual. As described in chapter 5.2 especially visitors with a technical engineering background focused on the cognitive part, instead of the emotional, experience-oriented approach. A first sketch of the surrounding is made and new or adapted expectations are formed. At this point the iterative cycle starts over again. In the case of an interactive sound installation the next stage is to perform an action producing sound. These actions can be accidental like the noise of steps or of dropping a jacket. Also initiated action are very likely to be performed. These actions include the usage of the offered objects like the whistle, the rubber duck, etc. in the room. The room responds to the action, data is collected, sorted and compared and a new interpretation of the environment evolves. The expectations are defined more clearly and eventually a visitor is able to predict the response to certain actions. At this stage also hypotheses are formulated to substantiate the overall picture of the environment. The conceptual model of the system is tested by planned actions, to verify the theory. Also inquisitive actions are performed in order to develop a deeper understanding. However, the interviews showed, that the group of engineers stopped their actions at a earlier point, once a sufficient explanation about the feedback was found. This group was found to be more likely to stop acting and therefore did not find further occurrences.

5.4 Evaluation results

5.4.1 Q1: What did you perceive when entering the sound installation at first?

The term perceive allows a wide range of interpretation and can be separated into all five human perception channels. Therefore the answers referred to the auditive as well as to the visual stimulus.

Visual stimulus: Six of eight individuals answered corresponding to the visual stimulus. Especially the cave, the loudspeakers and the cables hanging from the ceiling attracted the most attention. Only two probands mentioned the microphones.

Auditive stimulus: For two probands the first thing they noticed when entering the installation was the altered sound perception. One *professional*[21] explained that the system works with delays with different length. The second proband (*none professional*) referred to the enhanced reverberation of the room. After their first explanations all probands were asked if the auditory perception of the room was identical to its visual

21. member of group1: *professionals*

appearance. For all probands the room appeared to be much more reverberant and bigger than it was expected corresponding to the visual appearance.

Interaction threshold All interviews showed that when entering the installation the probands didn't know what to do or to expect. Only after an incident that produces sound the probands started to find out that the idea was to interact via acoustic signals. Therefore an *interaction threshold* can be defined. The *interaction threshold* separates on the one hand a none interactive behaviour of a installation visitor and on the other an interactive behaviour. Naturally, it depends on the proband how high the *interaction threshold* is set. It is also very likely that a visitor, who enters the installation alone and doesn't get any information, leaves without interacting or producing sound at all. This effect occurred several times, with visitors who were not involved in the evaluation.

For 4 (3pro/1non) probands the *interaction threshold* was very low and therefore the sound of their steps was enough to clearly understand that the room was changed in it's representation. At first, listening and recognizing known sounds was the main focus for these probands. They described the installation as dripstone cave and broad reverberant room with delays.

5.4.2 Q2: Did you produce any sound on purpose and what was the reaction?

All probands produced noises and sounds using:

(a) their voice

(b) their body (clapping hands, stomping feet,...) and

(c) the offered sound objects in the installation
   · rubber duck
   · snare
   · whistle/pipe
   · ...

Once the probands interacted with the installation all of them detected delays and echoes. Nevertheless, the perception of the delays was different according to the mainly focused aspects by the individual proband. The represented delays form a complex, multilayered sound scape where three focal aspects can be detected:

Specialization: A sound within the installation spreads over the whole room, forming single sound sources and not a continuous reverberation. Certain path of specialization were constructed by the delays. It was not possible to find a universal pattern, structure or law how these paths were constructed. Probands described it either as a circular movement in the room, coming towards the position of the sound source or moving away from the sound source, A third discription defined it as a randomly distributed process. Theories about the coherence of action and response included mainly the position of the sound source. Also a strong feeling occured for three probands that on the right hand side of the installation (snare position) the delay responses where enhanced and easier to percept.

Modifications: The delays are not exact copies of the produced sound, but altered or modified representations. The modification is mostly noticeable as a pitch shifting. This effect was mainly detected by singing or playing a melody with the pipe. The response was heard as a transposed melody that formed either a upwards or downwards movement. It was also described as a mirrored melody. One proband tried to produce triads with the pipe but the responses were too strongly varied. Nevertheless, a better understanding of the system's reactions was achieved, although the outcome was still unexpected. Another effect was the stretching and the shortening of the melody .

Sound pressure level (SPL) detection: Interacting with varied sound pressure level of the input signal shows different outcome of the represented delays. This fact was noticed by three probands but only vague information about the quality of the resulting sounds were given. One proband described that silent behaviour results in pleasant noises and loud sounds result in a loud and scary response. Another tried to force the system to become instable by producing sound as loud as possible and described the sound still as pleasing.

Single speaker listening: Only one proband listened closely to only one loudspeaker and described the delayed signals corresponding to this experience. The proband detected that every loudspeaker played back different signals, that form together the whole sound scape. By listening closely to one loudspeaker one could hear first echoes that were quite similar to the sound source, but after some time modified (distorted) echos appeared. The acoustic colour of the sound was more clear at a near position to one loudspeaker. In the room itself the changed spectra of the sounds weren't noticeable as strong as directly at the loudspeaker and also dependent on the frequency range.

5.4.3 Q3: Did you get any kind of feedback when producing sound?

All probands described the sound installation as interactive in the sense that action and reaction or sound production and response were coherent. It was clear that the personal actions were the source of the system's response. Yet, the sounds were at first not classifiable in the empirical knowledge of naturally heard scenarios and therefore it took some time to get used to the experience.

Problems occurred when two or more sound sources interacted with the system to distinguish between the responses. Nevertheless, probands felt it inspiring when two or more individuals interacted with the system. The sound was perceived differently when one is not at the position of sound production.

5.4.4 Q4: Have you been into the cave and what did you perceive?

All probands have been into the cave after some time spent in the installation. The questionnaire showed that the probands interpreted the sounds in the cave as either a recorded sound scape played back from a stored file or as signals captured in realtime. The second distinction was only made when a proband could find a connection between the sound and the park outside. Once a unambiguous noise was recognized it was clear that a microphone was set up outside the building. Predominant sounds where the noise of the fountain in the park and the noise of children playing, which was misleading for 3 to 4 probands and gave an impression of a scene at the beach. One proband thought the always present noise was recorded and the rest was in real time outside the park.

Interaction in the cave: No microphones were installed in the cave and therefore no interaction was possible there. Being in the cave one was forced into a passive presence. Nevertheless 2 probands found it very inviting to stay.

Sounds from outside the cave where not audibleor hard to distinguish for 5 probands, mainly because they focused an the happening inside the cave. For one proband it was clearly understandable but dampened. Another proband expected, because of this experience, played back sounds from other sources (ringing of church bells,...) on other positions in the room.

Open window effect: An *open window effect* is noticeable when sounds are played back in the cave and the scene happening in the park matches perfectly (e.g dog is standing in front of the window and barking). Although the position of the loudspeaker representing the signal and the position of the visual percepted source is not the same, the human brain merges these informations. The visual source position is predominant. [22] This effect was only noticed by one *professional*.

5.4.5 Q5: Did the sound installation change over time?

Overall the questionnaire showed that the system was perceived as a time independent sound installation. Yet, the personal perception changes while spending time in the installation space. As described earlier the *interaction threshold* has to be passed at some point, then experiences with on purpose produced sounds have to be made until one gets used to the scenarios.

5.4.6 Q6: Which attributes would you assign to the installation/the perceived room?

As mentioned before the predominant attributes of the perceived room were bigger and more reverberant. One proband realized another layer of the concept and explained that

22. verify with literature

because of the open structure of the room the outside is transformed into the inside. Therefore the choice of the room performance of the *Forum Stadtpark* was very good, considering this aspect.

5.5 Evaluation conclusion:

Eight probands with different backgrounds of their musical and technical training were interviewed about the sound installation. The questionaire showed that following aspects were mainly noticed by the probands:

1. enhanced acoutsic: bigger room, more reverberant, delays
2. coherent behaviour of interaction (once the *interaction threshold* is overcome)
3. specialization: delays forming paths
4. modification: pitch shifting
5. sound pressure level detection
6. time independent behaviour

Considering the three scenarios (i) feedback loop, (ii) reverberation and (iii) echoes/path which were designed during the creative process of programming, adjusting and listening to the developed system, the evaluation shows that scenario (ii) and (iii) where clearly percepted by the probands. Also the fact that with lower SPL inputs the system reacts with scenario (ii) and with higher SPL inputs with scenario (iii) is intelligible for the probands. Nevertheless, detailed questiones about the feedback loop scenario didn't result in a clear responses. Also showing actions like covering the loudspeaker membran to the probands gave no further information or hint about the existence of scenario (i). The parameterization of the feedback senario was chosen too subtil to detect single feedback frequencies and to destinguish scenario (i). Nevertheless, *Interstices* represented itself as a interactive, complex and multilayered sound installation, that was well received by the public.

Suggestions by probands: During the interviews someexpectations by the probands were recorded. Here these suggestions are presented very shortly:
· offer more objects to interact with
· objects with low frequencies (bass): all sound sources where either percussive and impulselike or consisted of high frequencies (pipe, rubberduck)
· played back sounds from other sources (ringing of church bells,...) on other positions

6 Conclusion

References

[CS90] Corbin and A. Strauss, "Grounded theory research: Procedures, canons, and evaluative criteria," *Qualitative Sociology*, no. 13, pp. 3–21, 1990.

[DK06] I. Dooley and L. Kale, "Quantifying the interference caused by subnormal floating-point values," in *Proceedings of the Workshop on Operating System Interference in High Performance Applications*, September 2006.

[dS05] L. de Soras, "Denormal numbers in floating point signal processing applications," April 2005.

[gra] Newton's law of universal gravitation. Wikipedia. [Online]. Available: http://en.wikipedia.org/wiki/Newton%27s_law_of_universal_gravitation

[Gri98] J. Gribbin, *Q is for Quantum: Particle Physics from A-Z*. Universities Press (India) Pvt. Limited, 1998. [Online]. Available: http://books.google.at/books?id=rS_8BUE7eN8C

[GS67] B. G. Glaser and A. L. Strauss, *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York, NY: Aldine de Gruyter, 1967.

[HBEM01] Y. Huang, J. Benesty, G. W. Elko, and R. M. Mersereati, "Real-time passive source localization: a practical linear-correction least-squares approach." *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 8, pp. 943–956, 2001. [Online]. Available: http://dblp.uni-trier.de/db/journals/taslp/taslp9.html#HuangBEM01

[Hen04] C. Henry, "PMPD: Physical modelling for Pure Data," in *Proc. Int. Computer Music Conf.* Coral Gables, Florida: International Computer Music Association, Nov. 2004. [Online]. Available: http://drpichon.free.fr/pmpd/

[Hof08] D. Hofer, *Implementierung eines 64 Kanal Mikrofonarrays*, 2008, master's thesis.

[Hor12] A. Horvath. (2012, July) Faster division and modulo operation - the power of two. online. [Online]. Available: http://blog.teamleadnet.com/2012/07/faster-division-and-modulo-operation.html

[Izh10] R. Izhaki, *Mixing Audio*. Focal Press, 2010.

[Moo90] F. R. Moore, *Elements of Computer Music*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1990.

[MP12] G. Marentakis and D. Pirrò, "Exploring sound and spatialization design on speaker arrays using physical modelling," in *Proceedings of the 9th Sound and Music Computing Conference, Copenhagen, Denmark, p.55-60*, 2012.

[NS14] H. Niedrig and M. Sternberg, *Das Ingenieurwissen: Physik*. Springer Berlin Heidelberg, 2014.

[Per14] (2014) Spatial signal processing (beamforming). Wikipedia. [Online]. Available: http://www.personal.psu.edu/faculty/m/x/mxm14/sonar/beamforming.pdf

[Smi07] J. O. Smith. (2007, September) Introduction to digital filters with audio applications,. Center for Computer Research in Music and Acoustics (CCRMA), Stanford University. [Online]. Available: https://ccrma.stanford.edu/~jos/filters/Direct_Form_II.html

[VMRL03] J.-M. Valin, F. Michaud, J. Rouat, and D. Létourneau, "Robust sound source localization using a microphone array on a mobile robot." in *IROS*. IEEE, 2003, pp. 1228–1233. [Online]. Available: http://dblp.uni-trier.de/db/conf/iros/iros2003.html#ValinMRL03

[Wei08] S. Weinzierl, Ed., *Handbuch der Audiotechnik*. Berlin and Heidelberg: Springer, 2008. [Online]. Available: http://www.worldcat.org/oclc/244008648

[Wik] Wikipedia. Sound installation. [Online]. Available: http://en.wikipedia.org/wiki/Sound_installation

[Wik14] ——. (2014) Side lobe. [Online]. Available: http://en.wikipedia.org/wiki/Side_lobe

[Zhi] L. Zhigilei. Mse 4270/6270: Introduction to atomistic simulations,. University of Virginia. [Online]. Available: https://www.koofers.com/files/notes-d0j5zhhir9/

[Zöl02] U. Zölzer, *DAFX-Digital Audio Effects*. John Wiley & Sons, ltd, 2002.

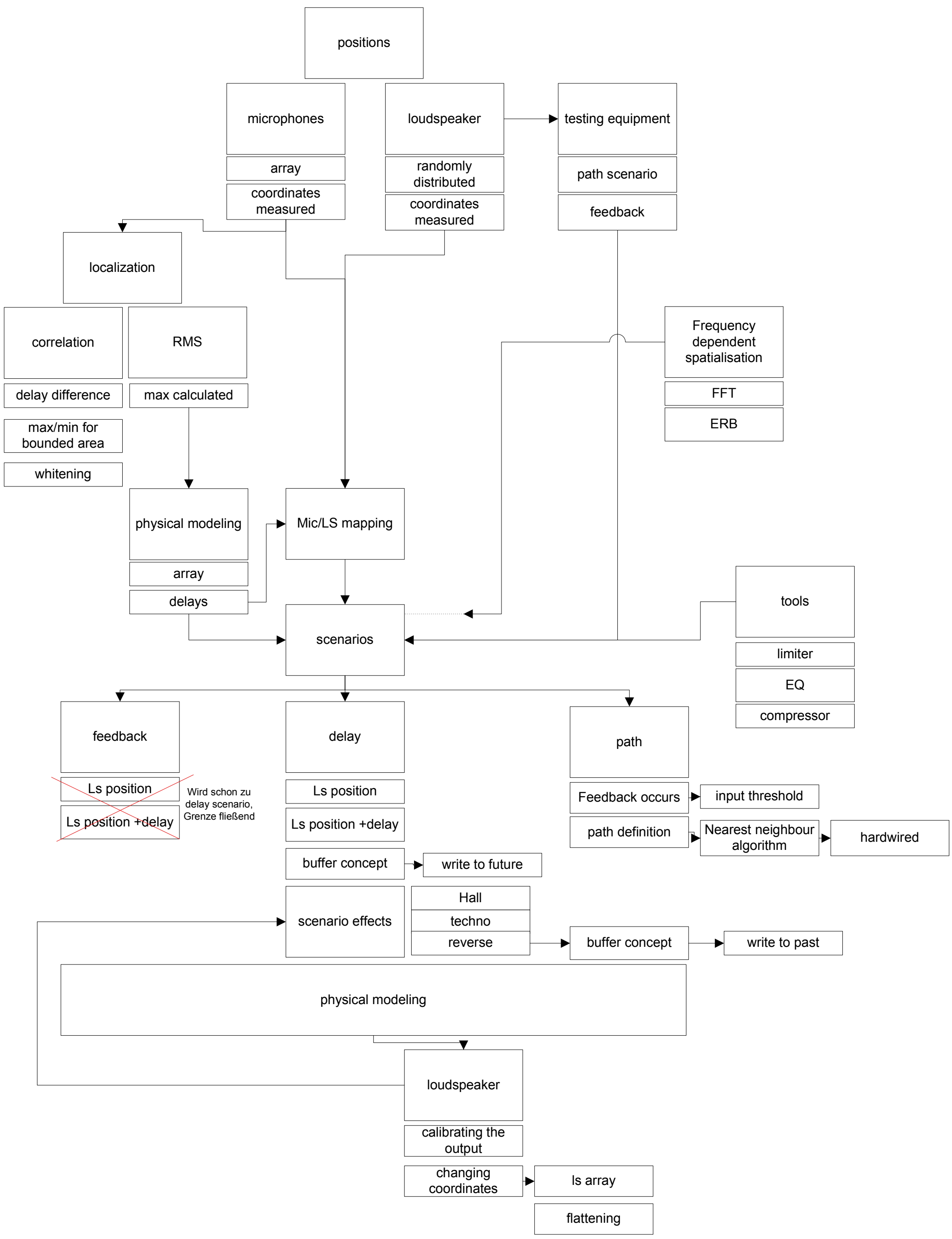


Figure 21: decisiontree.