An Audio Tutorial
Sightseeing sound

Bjørn Jacobsen
2012
Ma.IT. Audio Design and Ma.Musicae.Emusic
Ba.Musicae.Emusic

Content

Compression

Limiting

Gating

Phasing

Flanger

Reverb

Impulse Response

Noise

Delay

EQ

Vocoders

Field recording

Foley

Binaural recording and playback

Psychoacoustics

Synthesis

Modular synthesis

Granular synthesis

Additive synthesis

Game Audio

Game Engines

Audio Engines

Wwise

fMod

About the author

Bjørn Jacobsen, born 1981 in Copenhagen, Denmark.
Educated MA. of Music / Cand.Musicae – from the Royal Academy of Music in Aarhus, Denmark-Danish Institute of Electronic/Electroacoustic Music in Electronic music composition (2013)

MA. of IT Science / Cand.It.Audio design from Aarhus University and IT-University of Copenhagen (2013)

BA. in Electronic music composition and music teaching from the Royal Academy of Music / Danish Institute of Electronic Music (2011) – University Diploma of Game Audio Design (DADIU – Danish Academy of Digital and Interactive Entertainment – 2010)

Son of mother Susanne and father Bo, two both personally, creatively and professionally different personalities, they divorced in 1984 and have both new relationships and children, giving me two "half" brothers, Peter (1994) and Lasse (1995).

Brought up among fathers guitar and piano as well as mothers sewing kits and creative clothing designs helped grow my own creativity, but their professional lifes as a laboratory worker and later head of and a programmer was quite surreal.
Neither of them had made a living of their hobbies as of yet and in the late 80's and 90's Denmark, doing so was not directly frawned upon, but still not a very accepted approach to life.

Videogames where starting to become a regular part of family homes, including ours, or at least it did when I badly as a kid where asking for a game'n'watch for my birthday at the age of 3. As well a quite different musical taste compared to the other children at my age, up through kindergarten and later junior school, came almost out of nowhere. Yet my father listens to music which is not usually heard by the broad public, with a more geeky approach to music and my mothers listening to all kinds of music – for some reason inspired me to seek out new music experiences.

With the invention of cable tv and the introduction of British MTV in Denmark, in the late 80's and beginning of the 90's, music was finally introduced to me as something open minded and not closed as I had thought it to be previously, yet MTV had a wide selection of music playing, they had programs only for metal, only for techno, only for rap, and so on, before all of those genres became so pop influenced that they took over .

Up through the late 80's and early 90's, videogames started to take more and more of my time, and in 1985 I was given a Magnavox Odyssey II as a present, bought to me by my parents from a couple of friends they had, whose son, Peter, no longer used it.
In 1990, for my birthday, they had saved up for a commodore 128, which introduced me to a whole new world of games as well as programs for it that could sequence music. Not very good music, and at that time I was not a very musical person, listening to music and intersted in music, I had no clue on how to produce music or even given my own music a single thought.

Later in the 90's I began playing the guitar, under heavy influence of my father, yet I didn't bother much for learning it, as most of my time, at which i should have practiced was used on either playing videogames or my at that time hobby, building my own board games. Later my own games and the home tv videogame setups where exchanged with a varying amount of game consoles and computers. From the 8086 computer, to the Nintendo NES, over an amiga 500, upgrading the a 286, 386 and 486 – later pentium. Changed my life completely.
Even with a quote from my father and his at that time wife "you cannot make a living of making or playing videogames – you will end up being able to do nothing and just sit at home making no money and playing them" and "MUSIC! Not techno, nobody listens to it, you cannot make a living out of it" – I guess that was my queue card to prove them wrong.

In 1996, 14 years old - Danish High school was about to start the year after – and with that the guitar was kindly returned to my father, with a promise to continue playing music but just not the guitar way.
Earlier I had discovered that short circuiting the amplifier made wonderfull noises, the loop and distortion pedals could be used to make wonderful feedback, all which in my ears where wonderful music and not unbearable noise as well as having discovered fast tracker and other minor tracking programs for my amiga and begin looping everything to death and creating beats. I suddently catched a large interest in creating single sounds and spending time thinking of how I could design it or record it, making everything personalized – an interest of beats, compressed sounds and soundscapes developed.

In 1997 and up until 2000 where I graduated from High School, Oeregaard Gymanium in Hellerup, Denmark. I had discovered tools like, Fruity Loops, Rebirth, Cubase and hundreds of tiny other programs for my PC computer which could playback sounds.
Having said that as a positive thing, the downside of the teenage rage going crazy in sounds and making electronic music was that I believed it all sounded wonderful and trying to have at least one or two tracks played at every party that the class held – giving reason for quite heavy mocking from my classmates. But in the end it resulted in a completely packed school hall for the introduction party of the third and final year of high school, with me and my act "Cujorius One" playing live for one hour of goa,technoish, trance music.

The mp3 format had just been introduced a few years back and was considered a huge threat to the music industry at the time, I met a lot of electronic musicians who where both for and against it, but mainly the ones of the Copenhagen based acts of trance, minimal techno and so on, where against it, they believed that having your music copied would result in nothing but your act making no money from sales. Perhaps because the format didn't exist when they started out and they had to start with a different approach.

I released a few tracks online through mp3.com, an online music forum website where you could upload everything you wanted and others could hear it.
This was my first attempt to reach out to record labels, and in early 2000 Daniel Mondry of Berlin based psy-trance / techno act Zerotonine emailed me and said he had forwarded my information to their record label in Germany, who just weeks later contacted me and we made a deal of releasing some music and working together – so from here on, my electronic music career had been founded.

Through the many years of travelling as a musical act and dj'ing, my interest for the actual technical part behind it all started to grow, my lust for designing specific sounds and "producing" more than composing specific genres of music was still more and more appealing to me and i began to study what i could do with all this "skill" and interest. How could I make a living out of all this, if not just by playing music and being a DJ.?

I began a home study of composition and sound design, but found it very hard to begin and get the hang of it, mainly because I didn't know anyone in this field of craftmanship and I didn't know of any schools that could help me in any way.
I began searching for a book that could help me with all my questions, but never found one that really gave me the right answers, perhaps I wasn't looking in the right places, but very few people know where to look if not guided in a certain direction.

I began to apply for all the audio schools i could find, only to discover that most audio schools in Denmark requested that you had either proof of skill or where skillful in areas where I didn't feel skilled and was applying to the school to become so. I applied anyway, and applied to everything I could find, film school, the music conservatory / the academy of music, theater school as an audio technician, audio engineering school, acoustic studies with the university and so on and on, without much luck.

When schools won't let you in, you have to get to work yourself – that was my basic philosophy, and with that in hand I began to contact every possible musician, audio engineer, sound designer, film maker etc. Etc. To let me in.

Finally it paid off and I was allowed in on a small film set to function as an audio recordist on set, and I was contacted by a few people who needed some minor sound effects and field recording done for their productions, art installations, film post production and more. This was really it, I decided that this was no longer just a wet dream of a teenager and a young man, but serious business of what I wanted to work with for the rest of my life.
I immediately signed up for a three month audio engineering course with the nearest music school in Copenhagen.

In 2005 I got news of the Danish Academy of Music in Aarhus had started a new BA. And Ma. programme in Electronic music composition – which I applied for, but wasn't allowed in until 2008, for their BA. Programme of Electronic music composition which I finished in 2011 and then used as an access point for their MA. Programme of electronic music composition and an MA. Programme of audio design, game theory and more with the University of Aarhus. Both which I finished in 2013 – Including the diploma study of the Danish Academy of Digital and Interactive Entertainment, which i finished four times instead of the usual once.

During this period, along with the studies, I started to get more and more work with filmmakers and tv producers, making everything from small time short films no-budget, to low-budget to large budgets, greatly varying my pay, but it was worth it. Everyday was a learning experience and a joyful one. I got in touch with many A and AA game producers in Denmark during this period as well, making many different small games, some where only sounds needed to be produced and some where implementation had to be done by myself, some scripting for manipulation of the actual audio parts as well as scripting for implementation purposes.

All along all these periods, I discovered hundreds of little niche subjects of sound that I wanted to discover and research deeply into, but couldn't find help to study these and through my work in the field and with a "let's just go out and do it" attitude, I ended up learning a lot on many different subjects and paying attention when reading specific subjects and listening with people with knowledge in fields where my own knowledge was limited, is key to everything.

Introduction

This book is meant as a guideline and helpfull tool to anyone interested in sound as a creative craft, a guide that I myself would have liked to have during my many periods of study, confusion and creativity within the sound industry. Both a theoretic approach as well as a solid practical creative approach to many questions, as I have learned a lot through both, but playing only with theory leaves you without a clue in the practical world, and if you want to work with anything creative you need to know the theory to help you move the limit of what you can do and give you extra ideas. Use the theory to boost your creativity and learn the practicals to boost your creativity as well, don't stop where theory says you have to, use the practical experience to push the theories.

There are so many different ways to work and some even believe that there are correct and wrong ways to work. I will try to differ between right and wrong here, but also give my personal feedback on each of the situations, because I have been there and learned a lot from doing it my own way and then later being told why a different way was better or discovering that a mid way between the two worked best for me.

Sound is a craft, be it you only working in a studio as an engineer or as a designer of mystical sounds for videogames or if you are an audio post production worker, a so called, sound designer – giving films that final touch before they go out.
To make good sounds you have to know what you are dealing with, and you have to know what happens if certain things are done to a sound, why does a sound "sound" the way it does – as well as taking all your knowledge and just experiment and use the experimentation results later to be your inspiration to some out-of-this-world creatively built sounds.

Believing that every tool can get the job done, it's just a matter of the way you sqeeze it, is a key thought when creating sounds. A game engine or tool which allows you to implement sounds, is nothing next to your own thoughts on what you can do with this implementation process and how you can tweak it. Be creative and don't think anything is what it seems or sounds like.

The title of the book, is a bit of a challenge. An audio designers bible, an audio tutorial, game audio bible. Etc. It's not a bible (in the old sense of the word, THE book, with answers to everything, because the answer to everything is 42, if you ask deep thought (Douglas Adams). Forget about the title, and just read on instead.

This book contains everything I have learned over the years, and everything I think that every sound designer should know. I wrote it because I have been unable to find one like it, but thousands of books on film sound, but none one the creative side of sounds, game audio and sound design. Some of the theories, made by academic theorists are good, and if you know them, you can use them, but also have in mind that a theorist may come up with a fancy word for something happening in a sound, but if they have never sat in a studio and made a sound, mixed sounds together or worked in the practical field of sounds, then their theories should not be considered laws of the sound world, but more as guidelines to ways of thinking and nothing more.

Let's get started.

Sound in a nutshell

Sound is physically tiny vibrations or changes in atmospheric pressure, to hear this we absorb these frequencies with our ears and brain to interpret them into something that our mind and brain can understand. The human brain can understand and the human ear can pick up frequencies usually between 20 and 20.000 hz. With age the upper max frequency is usually lowered.

The most basic sound that exist, is the sine wave. A sound which is so pure in frequency that it only contains one, as well as being a sound which does not exist in the natural world.
One can basically build any sound in the world by adding sine waves together, but to create something as unique and complex as say, a human voice, hundreds of sine waves would have to be combined and constantly changed for us to be able to interpret this change into anything useful.

A simple sine wave curve, which is a basic mathematical function, is our way of mathematically interpreting an audio signal, it explains quite well how small atmospheric changes can be picked up by our ears.
A sine wave has a center, 0. and "circles" around this center between any two given values but in the case of sound we just use 1 and -1.

This is to describe the 0 as the standard atmospheric pressure and 1 and -1 as any given change in a higher or lower pressure, this also stating that if at a different atmospheric pressure than 1 (which is standard for air on ground at sea level, 20 degrees celcius) the sound you hear is different from when at other atmospheric pressures (which is quite normal). This doesn't always matter, but just having it in mind explains quite a bit of why sounds are different under water or high up in the air.

These atmospheric changes, are then picked up by our ears, where a tiny system of bones and skin changes the changes into nerve signals which can then be understood.
The ear conscists (very roughly) of the pinna, which is the only outer part of the ear which then leads to the ear tube which untimately ends in the eardrum, a tiny piece of skin covering the whole area.

From here the eardrum sends on the waves into three little bones, call the hammer, the anvil and the xxxxxxxxxxx, which again ends up in a tiny piece of skin, named the oval window. From there the waves are again picked up, but this time sent on into a long rolled tube with a middle piece called "the basilar membrane". This membrane is completely air tight with a sterile liquid covering it. This liquid transfers the sound waves and on the membrane is little haircells which picks up the frequencies and then, and only then, has the sound become a nervous signal which our bran can interpret and work with.

All this, just for a tiny sound.
This gives reason to believe that with so many changes, change in material etc. The sound can impossibly be that same inside our head as it was just outside our ear. Or is that really true?
Well, to some extent.

We cannot know for sure, to be very philosophical, that it is the same, but we know from recorded sounds today that microphones placed deep within our ears or just outside or ears or even connected directly to our eardrum, that the electric signal picked up from the microphone is almost similar to what we could record just outside the ear.
But the human ear is more than that, because at the beginning of time, the ear was probably not used for picking up language or detailed sounds like we want to hear today, most likely it was just a tool to warn our brains of the dangers to come, the same as our mouthes where probably not used to speak languages but only to be a weak signaling tool.

A very strange physical human phenomenon then appears, because the human voice conscists of many frequencies but the ones between 2 and 4000 hz are the ones considered needed for understanding what is actually said. The so called formant area of the human voice.
The pinna and ear tube of the human ear, for a strange reason enhances these frequencies, making it easier to understand what is being said, which leads to the discussion of which was used first? The voice to speak or the ear to hear speech.? A long philosophical discussion which I will leave out of here.

But this states that what we hear is of course not what is there raw in nature, but we hear a humanly naturally filtered version of it. Which then also prove that since no people are alike or built alike in their ears then nobody hears the same thing, yet we build up a common understanding of what is good and what is bad.

Sound is therefore a very abstract thing to understand, it all depends on wether you think of it physically or in the philosofical plane. Myself is of the belief that the philosofical plane is by far the most complex, but as with so many other aspects of philosofy and academic studies things can be studied to death and researched apart and still not come up with a clear conclusion expect that most of us think alike and what is really means is that the physical world is what matters (at least in the craftmanship business).

Therefore more on sound in the physical world:
Just as with frequencies, and a limit to which of those we can hear, there is a limit to how much volume we can take.
Volume is the power of how hard the sound hits, how loud it is, and so on. We can pick up a minimum volume of 20 µ Pa. (Micro Pascal) – which to start with don't make much sense unless you are completely geeky about remembering physics in high school.

Basically you can talk about volume in any form you want, but the used type of measuring is dB (Decibel) – which is a mathematical term for showing the difference between two numbers.
Any two numbers, you could measure the speed of your car in decibels, if you like.

In the mathematical world 3dB is double the number.
So if you have a car driving at 100 km/h then +3dB means 200 km/h. But in sound it's slightly different, because here we need to discuss many things and not just speed, and we need 4 times as much power to double the amplitude of our sound, which results in a sound being audible twice as loud. So in the sound world a sound needs to be increased 6 dB before it's twice as loud.

In the physical world, when we are discussing a sound moving through air (or any other material basically) – we are talking about dB SPL : which stands for Sound Pressure Level.
There are many other forms of dB measuring numbers, but SPL is what is used in open air, you may have heard of dBv or dBvu, dB FS and many more.
They are all used but in different ways, because like we just said, that dB is a mathematical way of telling the difference between two numbers, meaning that you must have two numbers to measure between. You cannot say that something is a certain dB without knowing what it's compared to.

So to make it short, 0dB SPL is no audible sound at all or just hearable, which is 20 µPa as also mentioned before. So if your sound is 40 µPa, then it's 3dB SPL.
Which is not very loud.

Normally a human can hear sounds up to 100 dB SPL. Above this level it will get painful at some point and at further levels your ears will break, eardrum burst, the anvil or the hammer bones will simply fall apart. Not very nice and don't try this at home.

Of course it is individual for each person where the threshold of pain is, but above 100 dB SPL is seldomly a good idea for the human ear.

Game sound came a long way

One of my main issues with sound, production and theories, is that sound and it's production is not a very old craft.
Surely it was used already back when filmmaking started around the year 1900, but the modern sound of films and games and the technology to develop this sound, was invented within my life.

This is surely not a problem, but that also means that sound is still a relatively new field of study and interest to many.
Especially within games, where I remember playing games in the 80 and 90's where sound was actually a luxury, and stereo sound was completely out of the picture, real-time reverb calculations and echo could not be done etc. Etc. Etc.

Yes it is many years ago that sounds started to have these features, but I still remember them starting to get used, so have in mind that we are not very advanced when it comes to technology yet, because believing such would be ignorant towards that the future will bring better sound or better technology, just as in the 19'th century when certain machines where believed to be the ultimate tool and would never be replaced or the invention of computers where it was believed that the entire world as we know it, would only require a few of such for all our calculation needs.

Sound came a long way already, as the first video games where created in the 1950's and 60's (Tennis for two and spacewar) had no sound. Pong in 1972, had sound. One of the first consoles which where made to have sound.
Problem is that audio designer / engineer Al Alcorn of the Atari Corporation back then said, that he did not have the parts or technology to actually make any of the sounds wanted by the game director, so the sounds of pong are actually made from enveloping the electrical current in the other parts of the circuitry.

Space invaders, Taito corporation (1978) – had better sounds, one of the first examples of different responsive sounds, when your spaceship would fire upon the alien hordes above you and the occasional U.F.O. Flying above you as well.

Mario. Nintendo (1985) – had music and sounds, synthesised high quality sound. One of the best examples of so called isomorphic sounds. Cartoony and what would be called "Mickey Mousing" of the sound, where even jumps, falling etc. Have sounds.

Doom. Id Software (1993) – one of the first examples of having sound effects as samples instead of using synthesis. Where the sounds, because of the limited memory and harddrive space of the 90's, would have to be very low quality.

So it's not until the beginning of the 90's, that sound started to be, in games that is, remotely close to what we have and use today. I was about 10 years old here, and I actually believed that I was born into a space age, where we had reached a technological high where making things better would be one of the most naive things to believe. That statement I later had to withdraw when my day purchased a pentium computer, where it quickly became obvious to me that there was much technology to come!

I was sort of in denial, that games and their technology, such as black and white screens, where a thing that only my parents had. Old school, nostalgic, and something I would never ever see. I had that thought put away, but recently I remembered that the game boy, with it's black and white monocrome screen, was released during my childhood. So black and white screens, where not a thing of the past, and it's actually not a thing used very long ago, if you look behind you, it's just around the corner.

In fact, I remember when technology allowed for games to be released on CD, and many games

where then re-released containing dialogue, still in horrific quality but still dialogue. Enough to blow the average video gamer backwards, and games like The 7th Guest, Under a Killing Moon, 11th hour, Sam & Max – Hit the road, Lands of Lore – the throne of chaos, where released with not only better graphics, but also "better" sound. In terms of dialogue.
Fact is that the sound, as with anything else technologically new, is that the quality is not very high, it's just impressive enough that it's even possible.

7th Guest. A 2CD game, (TWO!!!) - had real actor dialogue and greenscreen video recorded on to it, yet in retrospect, horrific quality of the sound. Very very low bitrates, yet the game came with the bonus of the CD's containing audio tracks, playable in an audio CD player. Where the sound was regular CD audio quality, the dialogue in the games where very low bit rates. So compression of files, to keep disc space low and keeping quality high, also came a long way.

1989, the year of the release, of the Sound Blaster from Creative Labs.
Not their first sound card, and not THE first sound card at all, but the first sound card to be released where it was commercially available. Roland and other companies had before released sound cards which where more expensive and lower quality usually.

1989, the year where sound was suddently possible.
Later came the sound blaster pro. Which could play stereo and sound blaster 16 and 32, which could playback in 16 and 32 bit quality. Before this I had an AdLib soundcard, which could only playback music and samples sounds where at a horrific quality, even though the playback happened through normal speakers - All this within my lifespan!
So do NOT, take stereo, high quality or anything else for granted, because it was invented and implemented just a second ago.

Theorists perspective

With any creative skill, which is a craft and done in the practical world, comes theorists, who have never worked with microphones, computers or mixers, but invent terms for you to use for analysis of your sounds. I believe that all these terms and thoughts of these terms are not only good to know, they are essential. But have in mind, that these are not laws on how things must work, there are good ideas and industry standard expressions on how things work.

You don't need to know the word diegetic, to make diegetic sounds, or knowing that non-diegetic is probably your background music if the source is not presented in the film or game. But knowing these terms help you understand what is going on, and helps you think creatively if used properly..

One of the main issues that I have come upon while studying sound, electronic music composition, sound design, game and film audio, is that game audio, is still such a new medium, that the terms invented by theorists 50 years ago for the academic worlds analysis on mediums and for the film industry are somewhat obsolete, yet the terms they created can still be used in game audio, it's not the same as with film audio and vice versa.

Many theorists have gotten it right, and their thoughts should be considered when creating sounds. R. Murray Schafer for one, and has written a brilliant book named "The Soundscape". This book is about the sounds that surround us, and basically a brilliant explanation to how the world sounds in a natural sense, and this leads to sound ecology, ambience and the manipulation of such.

Sound ecology for one, is a good discussion. Because should ambient sounds of any environment be manipulated? The human brain knows billions of sounds, and if it doesn't know a sound it instantly creates a link between what it knows and this new sound, creating an artificial image of what the sound source might be.

Sound ecology, or the green peace of sound environments, as some have come to call it, is the discussion between sounds in real life, real ambient recordings and the manipulation of these. Humans are very influenced by the sounds we decipher from mediums, and the discussion about sound ecology is if humans are beginning to only know the sounds of the sound designer.

Sound designers use birds as an indicator of morning, insects as an indicator of dusk, flies as indicator of something bad and rotten, cities have traffic noise, space ships have engine noise and explosions in space are gigantic, and wait.
Most of these are true and creates the right reference in our mind, but we have no clue of what space ships actually sound like, and from an sound ecologists perspective, that specific town you are looking at in your game or film, might not be the actual soundscape you are hearing. The soundscape you are hearing could sometimes be any other city, with traffic noise and all, which actually causes a distorted perspective for the viewer.

The difference between sound design, audio ecology and real audio comes into play here. Because if you are watching a film, your brain does not care if the sounds it is hearing are authentik to the visual source, only if the sound is "real" in that sense.

Footsteps, created by hands in corn flour, are definitely not authentic footsteps, but may act as real footsteps in your game and film. Plastic bags stepped on or moved around, serve as great real sounds for clothing, footsteps and other things.
The importance is if the sound is "real" to your visuals, then the source, authentic or not, is irrellevant.
Some ecologist believe that sounds should be authentic, as non authentic sources distort our minds and make people come to believe that the soundscape of the world is actually sounding like plastic bags. The tiny changes in a sound of a hand in corn flour vs. The actual sound of a foot walking in snow, is absorbed by the mind and ecologists consider this a major issue when it comes to our true sonic perspective on everything within our world.

Diegetic

and non-diegetic, is a term that everybody should be familiar with. It's not the most important term in the world, but it explains quite a bit of what is going on.
Of course over the years, when theorists have come across problems explaining specific sounds, they have invented new terms, which resemble other terms, which in the end is only adding confusion to the equation.

Diegetic is a term used whenever a sound is "on screen", it doesn't mean it can't be off screen, but if you as an audience knows the source of a sound then it is diegetic.
The easy explanation is that the footsteps of the character is diegetic and the background music is non-diegetic. But what if the source of the music is displayed? The character turning on the radio for example, leads us to know the source of the music being played, from the radio and therefore it's

diegetic.

But what if there is background music, to which, late in the progress of the film or game, you see a band playing this music, then it was basically diegetic all along, but was considered non-diegetic for great parts of the piece you just witnessed.
Michael Chion, created a term for this, acousmatic. A term used for when sounds appear as non-diegetic for periods of time but suddently become diegetic because the source is revealed. This is known as de-acousmatizing, and is basically a term used when the mystery of a sound source is revealed.

De-mystifying, as David Sonnenschein calls it in his brilliant book on sound design for cinema.

A voice over narrator is non-diegetic, but the intercom announcer at the airport is diegetic.

Hundreds of new terms, and variations of this diegesis term has come to life. And diegetic sounds are often refered to as story sounds, sounds that relate to the story as an image or perhaps a videogame without an image, like Papasangre.
A diegetic sound can be both on and off- screen. Often expressed as internal or external diegetic sounds, making the difference between the footsteps of the character and the soundscape of the city, or birds singing in the background but you don't actually see the birds.

Diegetic sounds, can also be defined as sounds that act and support your created designed soundscape. They are usually manipulated by the physical laws of sound, vs. Non-diegetic sounds to being completely in the front of the sound picture all the time.
The sound of footsteps, as we just agreed upon, is diegetic, but how far away are the footsteps? Of course your designed soundscape is a depiction of the realistic world but can only come close to realism, but will of course never really be.

When a sound is further away, it's volume is lower, it's high frequencies are lower, all of these factors imply when creating a "realistic" sounding footstep sequence. And only diegetic sounds will be affected by this, non-diegetic sounds will not.

The functions of sound

When creating a game soundscape, or a film soundscape for that matter, there are functions that need to be taken into consideration.

One of the key functions of audio within a game, as opposed to film, where the sound is linear and used to support the story / plot of the film. Game sound also needs to be informative and supportive of the players actions, unless of course you have a completely aestetic angle on this topic and decide not to have any informative, supportive, responsive or any other type of sound, then go ahead, that's your own choice.

I consider games, at some points, to be like movies, but with an elastic band to all the sounds, because you don't know the timecode of the events, and if you do then it's not really creating the interactive audio, which is what I think the focus should be on. For now at least.

Sound should support the environment of your scene, it should support immersion of the game and finally and of course the diegesis of the game world and the players actions. As for music, the music can be given the same names, as sounds, in terms of analysis – environmental, how the music supports the players perception of the game world.
Immersion, the music supports the players engagement in the game world, cognitive and mythical.
Diegesis, the music supports the narrative / plot of the game.

It's really hard to find material on these subjects, which is my first reason for writing this book, and basically all there is to find is whats written by all the film audio theorists, like Michael Chion, Bordwell & Thompson, Munday, Phillip Tagg etc. All of them briefly touching the subject of audio in other mediums than film, but then again, not really.

Your game may have

Events:

System events
Game events

Events that all need an indicator of their existence, gun shots, footsteps, starting, pausing the game and so on.
The use of diegetic vs. Non-diegetic sounds, to understand the difference between an informative sound from the game system vs. A sound signaling that your gun was just fired, and it's a powerful gun and not a small one. Also named responsive sounds, feedback sounds and more.

These are the informative parts of the soundscape, whenever you score points, start the game etc. Signal audio and responsive audio. The names basically speak for themselves, but signals are information carriers of events happening within in the system – and responsive audio on game events or player actions, are feedback from the game to the player about that the action just performed was received and performed either right or wrong. Confirming, disconfirming, questioning since some actions require more actions to be performed afterwards, and so on. Warnings, Notice etc.

Another term of trans-diegetic is also out there, as meaning sounds in the game, be it diegetic or non-diegetic, which connect the game universe with the players universe. This term is basically a description of anything which increases immersion between player and the game universe, because of the sound alone.

Time:

Historical period of the visuals and / or story line.
Time of day

A World war II game, would probably benefit from a certain aestetic touch on the sounds that indicate that this is 1944 and not 2050. as would the futuristic space adventure from 2050.

Time of day, is a classic indicator of what is going on.
Bird song may indicate anything from a beautiful morning to the horrific truth about the killings having taken place over night. Insects may tell you of dusk is coming in the desert, but also indicate that time has passed and the carcas on the floor has started decomposing.

As back in the microphones and recording chapters, where aestetics of recordings is mentioned, the time of the historical period could benefit greatly from this. If you are aiming at sounding like microphones and recordings did back then at least.
But having something sounding old, looking from the ecologist perspective, would be knowing what back then actually sounded like, which basically can't be re-created, but let's trash the ecologist theory for a second and try to come up with what would need to be done. Not just on World War II, but when looking at imagery from a certain period of time, listening to music from the same period, the style of people, everything from the period of time, an idea should already pop into your head – and this is the idea you need to re-create, because already at this point, you shouldn't be far away from what other people may, soundscape wise, connect with the period you are re-creating. Fantasy games are a bit of a challenge though – if you rely completely on this theory and method.

On time:
Michael Chion came up with some brilliant terms on film sound and time, which can be transfered straight to game audio. His term synchresis, as a mix of synchronism and synthesis, the mental connection between audio and image, over time.

Space:

Indoor vs. Outdoor.
Size
Location
Distance

Indoors, you would usually have reverberation (unless you are in an anechoic chamber – where you theoretically have none, but in practice would have very little still).
Outdoors, have different roll-off when it comes to frequencies over distance.

The size of the room, (the size of the outdoors perhaps? ;)
If a room is small it would have more or faster coming early reflections in the reverberation, if a room is big the reverberation becomes longer and the early reflections might be more blurry than in the smaller room.

Location, is the sound to your right or left? In front or behind you. Or even above or below you? A study into binaural sounds would do good here.

Distance. Over distance sound changes, usually only very close up sounds have the highest of frequencies, indicating that something is close. Where the human ear is trained from instinct to react on such as danger and very close and imminent danger.
Further away object, have a distinct roll-off in high frequencies, telling your ears that they are further away, taking these factors into consideration makes creation of a 3D soundscape much easier.

With the use of real-time 3D audio now and almost real-time convolution reverbs in games and the ability to code very specifically when and where such effects should occur, one should make a clear aestetic decision on how to use these effects, not at all, all the time, or just where you as a sound designer or audio director find it useful.

Position:

in 3D field.
Proximity.
Movement.

Just as with distance to objects, and the binaural effect on sounds because of their location in the room, the thoughts on wether your image or audio soundscape perspective is 3D or 2D should be taken into consideration.

Proximity, the proximity effect, is what will happen if something is close and up right in front of you.
The proximity effect of a microphone, having it close to your mouth vs. Having it further away, changes the sound completely, and a sound recorded very closely will take great effect of this, and will almost be impossible to make sound as if far away. Usefull to note when dubbing dialoque as well.

Movement, is the sound moving? Is it taking effect because of it's speed in the sense of doppler effect – the effect which happens when cars with open windows and loud music drives by, or fire trucks driving by, sounding like the sirens change pitch whenever they are getting closer and further away.

Objects:

Size.
Material.
Animate or inanimate.

What is the size of the object? Is it heavy or light?. And what is it made of.
A heavy material gun, wrapped in soft plastic sounds different than a light weight gun with no wrapping falling to the floor or when being picked up.
Is the object animated? Is it actually alive or just a cold dead object like the gun.?

Communication:

with other players.

Does intercom or team speak interfere with your soundscape of the game or have you made sure that the soundscape is compressed whenever the other players are speaking over the team speak intercom.?

Acoustics

When working in rooms, spaces, areas, etc. Your surroundings alter the sounds you hear a great deal. Nobody said that what you hear is what is actually played back, just like the way that the ear works by increasing and decreasing certain frequencies naturally.

When in a room, the distance between the wall, floor and ceiling, the dampening of the materials and so on, all work together in altering the sounds.

Most people have experienced to be in a church or other stone type building or anything with a dome, and heard the enormeous reverbs and delays that can occur, but also that you are able to hear people further away than you normally would – or from angles you didn't expect because their voices are guided around the room because of the materials and angles of these.

Many have also tried to listen to music and then walk around the room (the audio engineers most annoying habbit when visiting "normal" people) and hear the altering of the frequencies.
Because standing in corners the bass can be louder, when standing in the middle of a room you might hit the spot where all the frequencies phase out each other or simply where a certain frequency is just at a 0 in pressure and therefore cannot be heard, or at least becomes very subtle.

Acoustic "phenomenons", can also be used to our advantage. It is not all about dampening rooms and getting rid of frequencies, because sometimes we could be in need of increasing the volume of something.

Let's say, a subwoofer – which put in the middle of a room is not to much use.
A subwoofer is "omni directional", meaning that it sends out it's signal in all directions – this is because of the natural behaviour of the lower frequencies which it represents.
Placed in the "middle" of the air, sending out sound in all directions, is called 1/1 omni or 4π.

4π because it can be considered to playback in the four pie like areas of a bubble – representing the omni spreading playback.
So placing it on the floor will be 1/2 omni or 2π. - when cutting the amount of spread to half, we also increase the sound pressure level på 6dB. Meaning the a subwoofer placed on the floor actually plays twice as loud as it's actually doing, and half of the volume comes only from the acoustics of the floor causing the volume increase.

But what if this isn't loud enough? We could then move the subwoofer to a corner of our room and even further half the amount of flooring, compared to the standing on the floor only – which will again halve the amount of π. So it becomes 1π. Or 1/4 omni.
Increasing the sound pressure with further 6dB, now causing an 18 dB increase from what the subwoofer actually plays – basically doubling the value of your woofer, if you can only afford a small one, but need a bigger one, but have a great corner which you can place it in.

Standing waves and room modes

In a room, there is always a mode.
A room mode is considered the frequencies which have their playground there. Any room, no matter how anechoic it may be, has a mode. Maybe a very subtle one, but it's there.
Room modes should not be considered harmful or destructive, but more as a guideline as to why something may sound wrong in a certain room, or maybe even sound better in a certain room but like nothing when you get to another room and setup.

As written earlier, about sine waves and sound in general, sound frequencies are cycles of air pressure changes, constantly being over or under the normal air pressure.
And the amounf of change per. Second leads to our frequency. - and since that we know the speed of sound (or at least can argue a certain physical standard value for this) – we can then determine the wavelength of the frequency.

The wavelength is important to us, because it tells us when the sound/frequency starts and stops, and when it restarts it's cycle.
The speed of sound is considered to be 342 m/second. Some argue that the value 340 is the value to use, others that anything between 335 and 350 is fine. I will argue, that using the value 342,3 is best if this was a physical experiment, but sticking to integers is probably easier, so 342 m/second it is.

Now, it should be noted that this is when we are talking about air, as in free space here on earth at sealevel and 20 degrees celcius (which by now already has ruined the argument of the values because it's very uncommon to have such constant meterological values). Not up on a mountain, at an under the sea enchantment party, in an airplane, no, on plain ground, in a room, sealevel.
The reason for this, is that when altering the atmospheric pressure, the speed of sound changes, altering the wavelength - so under water where the pressure is much higher than one, the speed of sound is faster and in outer space with no air, the speed of sound is zero (0).
(so theoretically your voice is different when atop a mountain or in an airplane, than when standing on the ground).

The wavelength can be determined by a simple calculus.
Wavelength, λ = speed of sound, c / frequency, f.
So: $\lambda = c/f$

So given that the speed of sound is 342 m/second and we need to know the wavelength of 100 hz, we can then divide 342 with 100 and get 3,42 meters.
And what can we use this for then?

This frequency is now considered a standing wave, or the room mode.
But this only works in one dimensional rooms, from one wall to another parallel wall. (hint: building a room without parallel walls and no matching distances between wall, floor and ceiling will greatly reduce the amount of room modes annoying your setup).

A frequency of 100 hz, in a room with walls parallel, hard surfaces etc. Spaced 3,42 meters or half 1,71 meters apart will experience the phenomenon.
The sound itself is not standing still, but the name inclines this, because it's called a standing wave and will be experienced as such, because if you position yourself right in the middle of where the wave crosses 0 in air pressure, the wave, because of the room mode and wavelength – will always have the value of 0 and therefore this frequency cannot be heard in this area.

Which is why walking around a room can sometimes sound like that some frequencies are just not present at some points.

All this room mode talk, is to to give you an idea of how important it is to know about these, and why it explains why so many control rooms, studios and so on, setup in a living room or a bedroom can sound from bad to useless.
(not saying the home studios in a bedroom can't sound good, but this is a good explanation if this is the case).

Microphones and Transducers

Microphones and transducers? Really is the same thing. A transducer is a tool which creates electrical waves out of an analog input, which is basically the function of a microphone.

A microphone, in the standard sense as we know it, records audio being received by it, it does so by having a system which translates the analog input of the sound which was driven by the air to it. Meaning that it converts the small changes in air pressure, which we call sound, into a signal which can then be enhanced, recorded, played back etc.

There are many kinds of microphones, but among the most used ones are dynamic and condenser microphones. There are many other kinds, which we will get to shortly.
A dynamic microphone works through very basic electro magnetic theories, that when having a spunned wire around a metal core, you can create electro magnetism and move the core if you add voltage to the spunned wire, but if you move the metal core, the system works the other way around, instead of adding voltage to the system and moving the metal core, you move the metal core and create voltage which can then be sent to any system of preference.

A good example is that a dynamic microphone can actually be used as a speaker, and a passive speaker can be used as a microphone.
Setting a speaker in front of your kick drum to record the sound of the drum, is actually not such a bad idea.

A device which then controls the amount of movement in the metal core, like a small piece of paper which you yell at, will be able to pick up the slightest changes in air pressure and minimally move the metal core.
Here is is already given why such a signal must be amplified quite a bit before it can be played back and heard at a useful level again. The amount of voltage generated by the metal core is so little that it takes a lot of power to make it count, the basic difference between volts and amps.

This also means that any system which can generate a slight voltage, can be used as a microphone and this can be spoken of as a transducer.

A condenser microphone works slightly differently and requires a small amount of power to work, better known as phantom power. The power creates a small electrical field inside it's capsules, which when picks up changes in air pressure, will send a small signal back, which we then can use.
A general rule on condenser vs. Dynamic microphones, is that dynamic microphones are more prepared for a heavy load, but takes a lot more physical power to respond to sound and a condenser microphone needs very little power to reproduce it's electrical signal, this making dynamic microphones useful when you know the signal is of a certain power and perhaps if you already know that the noise from the source will be quite loud, condenser microphones on the other hand are much more sensitive the loudness, but pick up much more detail in the sound than the dynamic microphone – when close mic'ing and picking up signals from a kick drum and a hihat, the kick drum containing much more energy and that "omph" as we know it, not needing much help in getting loud. The hi-hat on the other hand has a lot of frequencies at once and we can benefit from the condenser microphones system to pick up more detail on these frequencies, than we would with a dynamic microphone.
Yet, it's always a matter of taste and there is no right from wrong when recording something, as long as it sounds the way you want and can justify the recording quality.

Any device, which can create currant can be used to create an electrical signal which can then be recorded
Taking a 9 volt battery and connecting it directly to a speaker, will result in the speaker being full blown out. Not destroyed or not working, but a firm click will sound and the speaker will be fully in it's maximum position. Change the poles and it will be in it's minimum position and you shouldn't be able to tell the difference between the two clicks.

This is also where accellerometers and piezo discs become very interesting.
Piezo discs work by being a disc of cobber with a bit of quartz crystal attached to it – quarts crystal is a very strong material and is often used in clocks, where a tiny bit of voltage is added and the crystal will resonate in a very precise clock, which in the clock is of course used to keep track of the seconds passed.
The keyword here is resonate, and if power is added enough to a piezo disc, the quartz and cobber disc will resonate back at the frequency mathing the size of the disc, creating a very loud almost unbearable square wave like sound. They are mostly used in your average home devices needing the make small beeps and bleeps whenever it needs your attention, like bed room alarm clocks, microwaves, smoke detectors and so on.

The other way around, just as the dynamic microphone, is much more fun, from a microphone perspective of course. Sodding a hot wire to the quartz and the ground to the cobber disc and connecting it to your amplifiers microphone input, will allow you to hear any vibration added to the disc. You can't talk into the disc and hear it – but you can place it on a table or window, and touch the surface of the table or window and listen to that.
By doing so, we have just created a contact microphone, which can now be attached to any surface and listen to it's vibrations, like the sound of a fence. The sound of a car, inside the metal parts, an insect walking on your table, put it inside a bag filled with water and drop it in your freezer and listen to the sound of water freezing. Or melting again, if you want to try a different approach, use some temperature resitant cables, and stuck it inside your favorite chocolate cake and you can hear it become crisp, stick it into the ant farm in the garden and hear it come alive.

Accellerometers are much more sensitive and doesn't require any prepping before ready to be used, and will give much higher quality recordings than the piezo disc, but a piezo disc, with cables and some electrical tape to protect it should cost no more than 10€ or 10$ for 10 or even more of them. So taking the price into consideration, you get pretty far with piezo discs and home sodding, besides sodding is fun and a learning experience in itself.

Characteristics

Microphones have many different characteristics, and the best way to explain it is that all microphones record differently and some may even be altered in their area of recording.
Most microphones record straight forward what's in front of them, but what about what is behind the microphone?. Some microphones, techniques and characteristics record more and / or less what's behind them or to their sides, this is called the microphone characteristic – the the microphone itself and the technique you decide to use, greatly expands the world of what is possible within recording.

A so called Omni characteristic, records everything around the microphone, no matter if it's coming from either side, front, behind or above the microphone.

A cardiod microphone, or as some call it, a kidney characteristic microphone, given it's name because it's recording pattern is in the shape of a kidney. A Cardiod microphone is what you would usually see on a vocalist or during an interview, where you do not care about the behind the persons involved or to the sides of them, for that matter. It records most of what is straight in front of it, but almost nothing of what is behind it and only slightly from the sides.

If you want even less from the sides and behind, you can use super cardiod or even club characterized microphones, recording only what's in front of you, basically leaving out all noise coming from behind the microphone. Easy to use on boom poles where you cannot get too close to the object or person you want to record, but still can just point straight at their face and record their voice. Quality on such a microphone differs a lot and distance to the source is crucial.

Another key element is that microphones where characteristics are very directional, like club, super cardiod and so on, are much more sensitive to touch, giving that "useless" muffled noisy sound when handled.
Very directionally characterized microphones are often easily recognised by their amount of holes, and many wonder why holes make less sound.

This is simple because of the way you physically direct the microphone, it is done by phasing out the sound coming from the directions that you don't want, and by adding holes behind the actual wanted direction, you can phase out the input from that direction, from the wanted signal.
A very simple depiction of how it works can be seen here:

Figure eight microphones, work just as the name says. It records to both sides, but nothing from the front and back. Can be used if only one microphone is avail able and you need to record equally from both people talking, or between two drums where either is important.
Figure eight microphone, is always an important part of the MS microphone setup, which will be described in the section on stereo setups.

Inside the mic

Now we have described many of the different types of characteristics, but characteristics is one thing, another is how and why they actually behave as they do.
Already here some explaining can cause confusion, because as audiowaves are usually shown, as sine waves and therefore only moves from left to right on a horizontal line and shows the changes in air pressure, this way of imagining sounds can be confusing as to, how an omni microphone works.

An omni microphone has a single layer, which may be manipulated by the air pressure. This layer moves the coil or changes the electrical currant within the capsule of the microphone, and if displayed as just a small membrane which is hit by the horizontally moving sine wave, as mentioned before, then making sense is kind of easy to do.
If the sine wave is up, then the pressure is high and the membrane, coil and everything else is pushed slightly in, if the airpressure is low and the sine wave at it's lowest point, then the membrane, coil, etc. Will be able to move out of it's position.
All this transformed into an electrical signal which we can then translate into sine waves again to record, amplify or whatever we want to do with it.

But what if you membrane is pointing left, and your sound is coming from the right, so the sine wave which is read from left to right, needs to move left and hit the back of the microphone first?

Then it can make it quite difficult to understand, from the imagery at least, how the microphone works.

Air pressure isn't excactly sine waves moving in the air, yes it can be depicted as the mathematical function sine, but only if we looked at air as going from left to right and at a very tiny space, which is never the case.
When looking at a recorded sinewave, you are basically just looking at a graph, interpretted from the input of a scientific measuring device, which tell your tiny differences in air pressure.
But air pressure is all around us and could also be understood as water or any other material for that matter and it's only at one excact spot, the one of the membrane and coil, which is read by our microphone and knowing the sound does not only move from left to right – we can then better understand that a pressure change, even if coming from behind, will still cause our omni microphone to respond.

It can also be named a pressure microphone. Meaning that any change in airpressure, no matter from which direction it comes, the membrane will be affected by it.
The pressure at that excact spot, is "always" the same, no matter which way you turn the membrane.

Now another type of microphone characteristic and membrane setup is pressure gradient. A pressure gradient microphone allows the membrane to be manipulated by airpressure not only from it's front, but also from it's back.
Doing so will cause some weird, but nice and usable effects. Pressure gradient basically gives you a figure eight microphone, because the electrical currant produced by the membrane and coil will be the sum of the values coming from the air pressure from the front and the back combined – and if a sound comes from the side of the membrane, the pressure will be equal on both the front and the back of the membrane, causing silence – in your recording at least.

Having a regular box, with a membrane, which works as our omni microphone, making a small hole in the back of the box causes sound to enter the system from behind.
But if we dampen the signal fron the back of the microphone a bit, we can combine the two pressure and pressure gradient techniques and use the signal coming from behind as a tool.
Which will result in a super cardioid characteristic, meaning that it will respond to sounds coming from the front and in the angles around here, but not from the side, yet still anything coming from directly behind the microphone is still picked up.

But what if we phase invert the signals coming from behind? Then we are able to record what is coming from the front, but phase out the past of the signal coming from behind, making it possible to record only what is coming right in front of the microphone and not from the sides or behind – you may have seen so called shotgun microphones or zoom microphones, long, with a high number of holes or slots along the way down the microphone away from the main membrane at the end.

Having more holes in the back of the microphone, as you may have seen on many zoom and shotgun microphones, they look as if they are "good" in the sense that they have many holes and are then able to pick up a lot of sound.
What all these holes basically do, is that they allow sound to enter from behind and will phase out this signal from the original one. Meaning that the signal actually picked up by the main membrane is actually omni and everything around it, but by phasing out the signal from behind this then "disappears"

A common rule is that, the more directionally sensitive a microphone is, the more sensitive it is to your touch as well, so a directional microphone should always be used with gyro, elastic band, boom poles or steady mounting to objects, and not held in bare hand or touched too much during recording. And then again, don't touch your microphones during recording at all anyway.

Stereo Setups

When using two microphones, to record both left and right, we have a stereo setup.
It's not the same as recording two sources in mono, but it makes a great difference to your stereo perspective when listening to the playback or just monitoring the sound straight from the microphone in headphones.

An AB Setup, is a setup where two omni characterised microphones are spaced with about 18 cm aparts. This type of recording gives a very detailed and almost binaurally precise playback, it can be used to create a stereo perspective of a human listener, by being placed in front of a live band.

ORTF. L'office de radio est television France. As the name inclines, then this setup comes from France, and was used in their radio and tv productions. An ORTF setup is two cardiod microphones, put with their bottom closely together and their fronts with about 18 cm apart. Giving an angle op about 110 degrees at the bottom. This gives a wider stereo perspective than the AB setup, because of the different type of microphone and the way they point.

NOS. Nederlands Operation Setup. Is from Holland, and almost resembles the ORTF setup, but is wider setup, with more angle and further apart.

A question arises quickly when going through all these different, yet almost similar setups, of why?, why all these different setups when all they do is record and of course sound differently, by why so many different ones.?

In the past, when internet and other quickly distributable media wasn't around, in the pre-historic ancient days, almost back to the time of big bang, all media producing companies within the industry had to use a standard setup, this is to make it easier for everybody to use the same, so that when the radio and television company in France orders something recorded in stereo, the audio engineer won't have to wonder what type of setup they want, he just has to use the standard setup, which in this case would be ORTF.

Of course in these days, we can, as well as they could, use whatever we want. So try them all and get to know their results and use it to your advantage when making stereo recordings.

The XY setup, is almost a reversed ORTF. Where the front of the microphones are crossed, so that the right and left side are almost in the same location but right is turned slightly left and left is turned slightly right. Giving a very narrow but solid stereo perspective.

The MS setup, is quite different from the others. You will place a cardiod microphone pointing straight forward, and a figure eight behind it, for recording the sides. This results in two, not very stereo like signals.
But taking the figure eight recording, splitting it up into two signals one left and one right and phase inverting either one of them, will leave you with three signals. A cardiod which is dead center, and a

new stereo / phase interved two channel mix of the figure eight microphone. Combining these two signals provides a huge stereo perspective. Turning up the center microphone and to your own desire adding the stereo figure eight recording works really well.

You can also make a dobbel MS setup, which is two cardiod microphones instead of a figure eight, but the microphones must of course be alike and should be matched, therefore a set of matches microphones is recommended and dobbel MS setups can be purchased from many microphone producing companies.

Recording

It is an important and popular, yet misunderstood fact that recording must be done with high quality equipment to be good or usable.
Just as with mixing, clean mixing and knowing your EQ and compressors is a good start, but nobody ever said that breaking the "rules" where not allowed.

Understanding what is good recording practice never harmed anyone, but only when knowing good practice can one take advantage of it and twist it.

Ben Burtt, one of the greatest sound designers of all time, made it clear to me, when he was telling about how he created certain sounds.
Before that I believed in authentic audio and it was just a matter of using the right equipment, which would then lead to good authentic recordings which could then be used with just as convincing results and things that where manipulated to death and not authentic where basically cheating and not real. I guess I was a bit of an audio ecologist then, by accident, because I certainly do not agree now.

Ben Burtt, has shown that spaceships are microphones out the windows of cars, lasers are electric cables which short circuit, smashed tv's, recorded, re-recorded, filtered, re-recorded and then added some magic.
Realise that as a sound designer, knowing the equipment and how to use it, allows you to know what your recording might sound like and by knowing that, or thinking of that to begin with, will allow for great sounds to be created.

In the chapter on Rassool and A Mothers Inferno, you can see some of my personal examples of sounds created for small student game productions, during two of my three DADIU productions (Danish Academy of Digital and Interactive Entertainment).

Today, digital recordings are storing the data given the system, because of the frequencies and waves caused by the sound in the physical world.
A translater is needed, a converter. whenever a microphone or other type of transducer is hit by a sound or wave, the signal goes from the physical world into the electrical world.
Recording this electrical signal allows us to store it and re-play it. With modern computers we can then twist and turn it to death of countless manipulation..

A digital recording, converts the electrical signal into information about where the waveform is right now, which then allows for storage and playback as well.

There are many discussions on where a recording should be placed in the digital spectrum, and

some believe that taking use of the full bit depth and recording at max volume is the rigth way to do it, making sure it doesn't distort of course. Going above 0dB FS..
Some believe that when recording at 16 bits, this is necessary, but when recording at 24 bits, you should record at an average of -20dB FS. This makes sure that you have plenty of headroom before distorting and you still get the high quality of the 16 bits.

I believe that we can then argue and probably agree upon, that as high quality a recording as possible, is the way to go.
Not saying the 32 bits and 96 or more kHz should be the way forward, but recording in formats like that, makes it easier not to distort, get more quality in terms of frequencies and tiny details in your recording, and you can always after the recording format the sound to a lower format, which would be pointless the other way around.

Personally I prefer to record at 24 bits and 96 kHz, because I believe that such a high quality recording, when manipulated, such as being stretched, lowered in pitch, hightened in pitch etc. Etc. Makes better results, and then saving it in the format needed for your game engine or 48 kHz and 24 bits for film post production, is a good and healthy approach.

I not only believe it, I can hear it as well.
And using your ears is of course as always the best debugging tool we have.

But aestetically there is no need for all this, because if you are seeking an authentic sound of the old school – where low quality is needed, then recording at high quality and trying to reach the sound of low quality may be more difficult than at first glance.
Recording at old tape recorders, having magnetic tapes stored in moist surroundings for years before playing them back again, may have an aestetic purpose and may give you just the sound you need which the high "quality" digital recording might not have given you.

Don't be fooled by high quality microphones, they only sound good if used properly, a 10.000 dollar microphone may be equally crap compared to the cheap 1 dollar microphone you found in your cereal package one morning or the microphone in your phone, can be of great use if you need this type of quality or audio perspective, instead of as said before, recording at high quality and then reducing the quality.
Of course if you are not sure, then the high quality recording is better, because it can be manipulated which the low quality one cannot as much, so knowing the sound aestetics of your recording devices is a good starting point.

When recording, take into consideration which microphone you are using, as just written.
But also the characteristics of the microphone, does it have a low cut? Or a padding option, which allows for louder sounds to be recorded without distorting the microphone or digital system itself.?

When recording a certain sound, think of if this is really the sound you want.
The sound of a jet taking off, may sound great, but not depict the type of jet taking off sound you are seeking – so think of different recordings, that may sound similar but yet very different.

One of my many projects throughout my career, was my Ba. Project of electronic music composition. I decided to find some aestetically pleasant sounds within the human body when being manipulated, as well as in animals.

I did some quite weird things to get these sounds, but the results where beautiful. But the first thing taken into consideration, was of course, how do I actually record these sounds?.
I swallowed a DPA 4060 microphone, hooked up to a Sound Devices 788T and took a ride on the bus and the metro here in Copenhagen. The results are actually the sound of the bus and metro, through the filter of my own human tissue. The sounds are very drone like and muffled, but yet they contain sounds from all over the spectrum and after the recordings, I decided to experiment with their pitch, reversing them and the results where some really nice quality drones.

I created a stereo microphone out of a pair of iPod headphones, putting them together with some plastic kitchen wrapping and stuck it up my rectum, and connected the headphones to the input of my Zoom portable recorder, and took a walk on the ferry between Aarhus and Sjællands Odde in Denmark, a 1 hour out ferry trip recording the sounds of the ferry from my inside.
Most of the recording is useless, but some of it reveals some very very beautiful sounding drones, some which I later used in some of my productions.

I wore a pair of Sennheiser HD-25, and set them into the microphone input of my soundcard and went to the dentist.

I created a small piezo disc "microphone", and with some gum attached it to the upper part of my mouth, and then pressed record and had the dentist do his work.
Creating some really high pitched sounds and some weird noisy sounds from the touch of the gums.

Projects

Setting up tools and working in various environments is key to letting go of the box that so many stick to working within.
The box is never "the problem", but it can quickly become the part of any project or any creative mind that limits your productive powers to be either null or boring.

Here are a few projects, which will show how to work in many different environments, from making weird sounds out of nothing, manipulating existing sounds with various setups and creating your own programs, scripts or patches that will do good for your sounds.

Project one:
Setting up a minor dj tool in max/MSP using your laptop or desktop keyboard as a controller.

Through my many years as a professional electronic musician and DJ, i spent a lot of time playing live by either just clicking play and having a basic fail safe set or just letting go of the controls and don't really touch anything during the live set.
This got so incredibly boring that i decided to create some tools for me to use when i was out playing, even though there where professional tools already that could do these things for me, they always contained more than I needed and the building process of such a tool was destined to be a learning and meaningful experience.

I decided to build a DJ tool, which basic functions was to play two or more tracks at the same time,

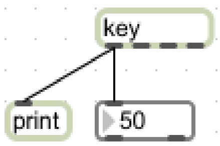
which could be pitched and manually synced / beatmatched.
In this way I would have a tool that was my own, my own way of playing my tracks and I wasn't using the regular DJ setup which was usually frowned upon by DJ's and other live acts, as a non live act thing to do.

When building a tool or coding anything, it's important to have an idea before you get started, and / or have a slight idea problem solving process. If you start with nothing and just go into blindness, you will sooner or later experience that your code or patch is super messy and takes more time to clean up and keep track of than actually working with it.
Therefore it's very important that it's clear what we will create, why and what it should do in the end and what not to do. A homemade piece of software, or any professonal piece of software for that matter should only function as intended, of course anything can be used as not intended often with cool benefits in performance, but this makes any kind of code / program buggy – and if this tool is to work in a live situation, not crash and so on, buggy is not acceptable (as if it ever was).

Here we will build this tool and along this way expand it on the go to show how already working parts of a piece of software can be expanded easily.
For this we will use max/MSP, which can be aquired at www.cycling74.com – a very usefull GUI based tool, if you don't have or want to use max/MSP, you can always use PD, Pure Data, which can be aquired at Puredata.org, but PD use slightly different, though free.
Max/MSP is also a newly integrated part of abletons LIVE software, so it will never harm to get into max/MSP.

First we need to open max/MSP. If you don't know max/MSP before hand, you should turn to the chapter on max/MSP first, or if you like, use it as a learning by doing experience by following this guide.

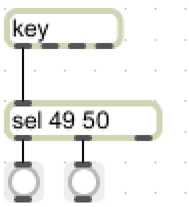
In max/MSP, create a new open patch. A blank piece of paper.
Press N for new object or if you like, double click (max 5 or later) – to get the list of colorful icons up to choose from, and create an object named key.
This object has several outputs and one is ascii number, which you can connect to the object print and here quite easily see the ascii number for any key pressed on your keyboard.
So routing the ascii signal from any key pressed can now be used as buttons to toggle anything you like.



Object	Message
print	49
print	50

Here you can see, that by pressing the keys >1< and >2<, we get the responce 49 and 50.

Using the object *sel* or *select*, followed by the ascii number, will route the key pressed out through a specific output, instead of just starting everything at the same time.
So connecting this to something which will play, will of course cause it to play.



Now we would like something to interact with and setup a buffer~ object and a groove~ object. The ~ meaning that the object is treating audio or audiofiles as in and / or output instead of just number / midi messaging. The buffer can store an audio file, and the groove~ will playback what is in the buffer, if they are named the same. So let's name them buffer~ track1 2 and groove~ track1 2. (Track1 is the name of the buffer, and the number 2 after the name is the amount of audio channels in the buffer – as well in the plaback object groove~ if we are certain that the music which we input have two channels / stereo / dual-mono etc.) Most objects can be controlled by so-called messages sent to them, and these are easily made by pressing >>m<<, and type whatever message you like, or you can hoover the mouse over the input of an object and get a list of inputs which the object can receive, in this case the message >>replace<< sent into a buffer~ object, will replace the current audio file stashed inside it with a new one. Opening a window in which to choose the file to open.

Replace changes the file loaded in buffer~

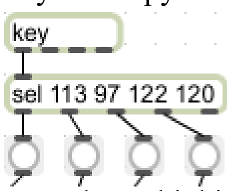
Groove~ needs a control signal to play, and straight playback is the number 1, but altering this number will cause in a pitch change. 2 meaning double the playback speed and 0.5 meaning half the playback speed..

This is a signal code, being sent at the speed of the sampling rate of your sound, usually 44,1 kHz or 48 kHz. So if the input is 1 then it will be 44,1 kHz * 1, giving the same playback as intended originally in the audio file. Max/MSP works out by itself, if the number 2 is sent and then doubling the playback speed, even though there are not double as many samples in the file. The number 1 must be sent to the object sig~ and then into groove~. The sig~ object just repeats the number given at the current samplerate as mentioned before.

The number 1 sent from sig~ to groove~ results in playback at normal speed.

Instead of just sending a message with our number 1 into the sig~ object, using a variable number as a number box, float or integer box is more flexible. The float object for more precision. Remember, you can always press ALT or command (on a mac) – and double click any object to get a very useful reference patch opened, which will quite well describe what any object does if used as intended.

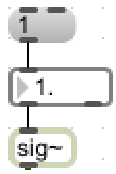
So creating a float box by pressing >>F<<, and routing the signal of buttons >>Q<< and >>A<< into adding or substracting a given value from the original number, we now have a controlled pitch value. And letting >>Z<< and >>X<< control start and stop, by letting X drop the pitch value to 0 and Z put the value back to the initial number. Try it out by creating a key object and find the keycodes of these keys, or do it the easy way and copy it from the figure here:



The sel object routing out the signal, when keys Q,A,X and Z is pressed.

Q pitches up, A Pitches down, X stops playback and Z resumes. If we route it in this order.
To do this, we need to be able to add a certain value to the current pitch, so we need to store the original number and add something to it. The reason for storing the number is of many reasons and one of them is that if we need to pause the playback, then we can return to the pitch we came from instead of returning to normal playback pitch and having to manually remember at which pitch we where, just two seconds before having to go live and into the mix.

Storing a value, can be done with the object float. Not to be mistaken for a float value object.
Let's create a message, which starts normal playback, by making a message which sends the value 1 into our number box controlling the sig~ object.

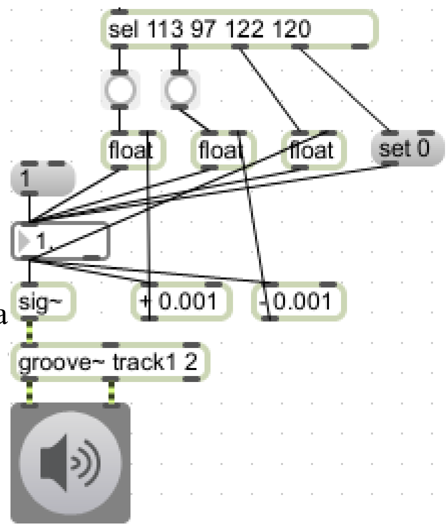


This message of 1, can now be clicked with the mouse when the patch is locked, to start playback – it can be linked to any input if you like.

Now, we shall add and substract a value of 0.001 to the value sent to the sig~ object. We do this by sending the original number in a + 0.001 object's left inlet. Note than most objects also do the math or functions, whenever information is received. + 0.001, means that whatever we send into the left inlet is instantly added this value. - 0.001 would of course substract the value from the number. So making both objects already, wil give us the possibility to both increase and decrease our pitch control number. Sending the result of the new float value to a float objects right inlet, will store the value inside the float object. If the left inlet receives a signal, from our pressed key Q or A, the value will be sent out through it's output, which we can then route to our pitch control value. In this way we make sure that our value is stored and ready to be used whenever we need it, but will never stack or overflow, or even be called, if we do not press the key.

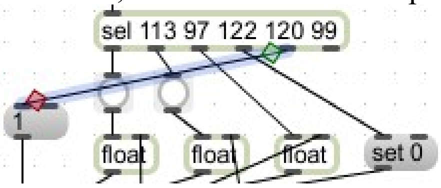
Now connecting our sel output, from the key inputs into the float object, we can now control this. Adding an extra output from our pitch control number into a third float object will let us store the current value instead of a slightly increased or decreased value, this will allow us to add the functions of the extra two keys Z and X into the patch but connection Z to that float objects left inlet and X to a message >>set 0<<, set meaning that the number 0 is sent to the object but no further, a great way to control the signal path. This will allow our playback to stop whenever X is pressed and to resume whenever Z is pressed. Their values are 122 and 120.

Already now a patch can look a bit messy, so encapsulating some of the objects is a good idea, also aligning or rerouting the patch cords can be done by pressing ctrl-shift-A or Y.



This allows us to playback the audio inside our buffer and we can increase and decrease the pitch, pause and resume playnack, just by using these four keys. Q,A,Z and X.

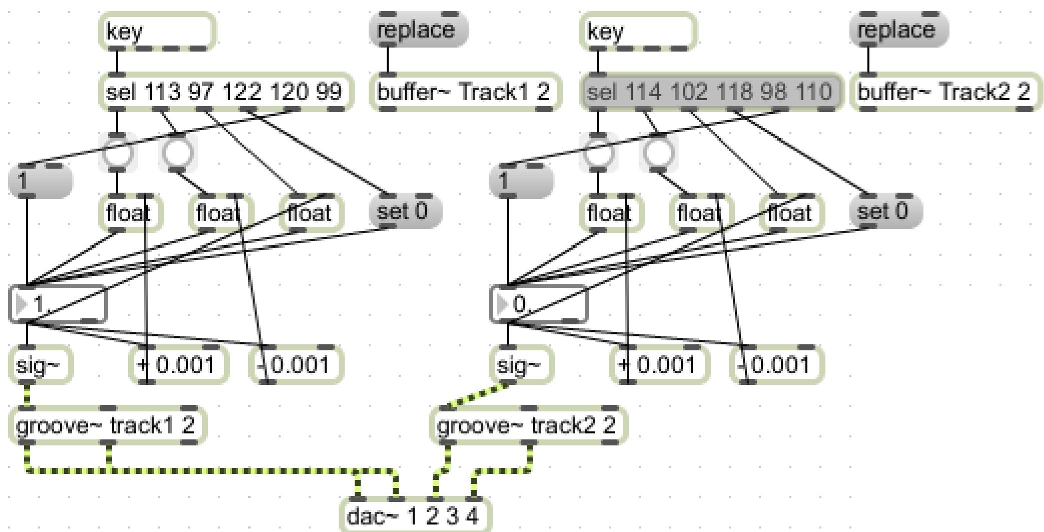
Basically all we need to do now, is dublicate the patch and make sure that our key input is different for this one, and that our buffer~ and groove~ objects are named track2 2 in their extension instead of track1.
But the little message of 1 which is out there to the left, could become of use, because what if you are already pitched up or down and you don't have time to press the Q or A key 25 times, then having a key which resets the pitch, might be a good idea. A little help from an experienced player, would be to place the reset key away from the other pitch keys, so let's choose key C for this.
Ascii value of C is 99, so sel 99 and dragging from the newly created outlet of our sel object over to the message box containing our number 1, we can now reset the pitch.



The setup works fine so far, but when we add track2 to the patch, we shall need more outputs, taking that you have more than just a stereo out on your computer of course.

Our current output is the grey icon with two inlets. This object is actually named ezdac~ and doesn't provide any ability to output more than stereo.
There is another output, dac~ (which stands for Digital to Analog Conversion – simply put from digital computer to real world audio)

this object can be added the extension of as many as you like outputs, let's just name them 1 2 3 and 4. so dac~ 1 2 3 4, gives us an object with four audio inlets, which internally in max/MSP can be routed to whatever output you desire, in our case probably the first four outputs of your soundcard. So let's connect the two outputs of track1 to the first two inlets instead. And our track2 to the third and fourth inlet. Creating a new key and sel object to control the control inputs of track2 can be a good idea, but they could also just be added to the other key and sel object, but to make it less visually confusing we'll do it the other way here. And use keys next to the other set, so R,F,V,B and N.



So now we have a fully functional little DJ tool, without too many gadgets and gizmos connected to it. Connect the four outputs of your soundcard and cue up some music in the track1 and track2 buffers. Play them and mix them, any way you like.

But playing on your laptop, with the keyboard can get alittle messy, and a little annoying at some point when the guys up on the stage do nothing but stare into the laptop to make sure that they press the right buttons.

And we aren't really done with adding gizmos.

Project one midi expanded: Using any kind of midi input to expand your patch, using a keyboard or any key or knob based midi controller.

The standard setup that I used to use, was my Clavia Nord Modular 2oct Keyboard, this was my very first attempt to make this tool, and I used it for many live acts of mine. Though the keyboard model, though interesting and fun to work with proved unstable after a while. Not by crashing, but proving that there is a reason why two turntables and a mixing desk is easier to DJ with, than pressing keyboard buttons.

One of my ideas for to enable the use of samples, so that the keyboard itself could get a more "musical" function, than just to trigger and pitch the music.

Project one expanded: Using a Korg nanoKontrol2 as a controller.

So, I went out and bought a Korg nanoKontrol2, because I had an upcoming live gig in Tel Aviv, Israel and decided to hook it up.

My previous attempts to route out the information from a wireless numpad and a Clavia Nord Modular 2oct keyboard showed really good results – but both of them didn't have faders, which resulted in my setup not to be using such, and a professional controller like the nanoKontrol2 would provide me with all I needed, and it's very flexible as well in terms of which controller goes to which knob and fader.

Project two:
Setting up a vocoder to manipulate a voice

Setting up a vocoder to manipulate other sounds

Setting up a vocoder to mess around

Project three:
Coding a small skipping sound in c# using unity 3D

Coding a dynamic audio tool in c# using unity 3D

Coding a system to minimize the amount of code for every sound in c# using unity 3D

Project four:
Have your webcam make sounds in max/MSP

Have your webcam change your sounds, by using fiducials

Excersise one:
Spectral listening
Spectral listening is a key tool and exersise to getting into frequencies and how they work in the auditory field. Yet one can always argue that you can just fiddle around with an EQ or other tools and eventually it will work and sound right, then we can also argue that getting to know frequencies and how they actually sound in a hi-fi environment can never harm.

When knowing the different frequencies and my getting used to spectral listening of sound environments, then you will noticed an increase in how fast you will get it right, instead of all this fumbling in the dark, finding that hum frequency (if it's not just 50 hz from your power supply).

Spectral listening, the exersise at least, is done by listening to pink noise and add or remove frequency changes to this noisy spectrum and then be able to tell which frequency, Q and so on was changed.

Try it out. Put on some lovely pink noise, not too loud and not too subtle, and add a regular 6 dB/octave EQ filter on it. Now choose a random frequency, and add 6dB or 12 dB to it.
Of course, this frequency becomes more present, but try doing so with other frequencies and then hum with what you hear, and / or even give what you hear a name.

A great example, is to make a list of all the frequencies. I have made one for you and placed it in the back of this book or you can get it from my website.
Try going through all the frequencies in octave like chunks, and give them names or at least, write down the tone you hear.
32, 64, 128, 250, 500, 1000, 2000, 4000, 8000, 16000 hz. - What you name your frequencies is out of my hands, but it's obvious that the 32 hz increase in frequency will be lower than 1000 hz and so on. Some like to name their frequencies after letters, like AAAAAA or OOOOO, and then AEAEAEAE and OUOUOUOU – some like to name 8000 hz, the gardening water hose, and the 16000 hz the hissing of the dead. 250 hz is a missile, 500 is a waterfall etc etc.
Find your own names, letters, numbers, whatever. Which makes you remember the frequencies.

Now this drill is easier to do if you are two, or an entire class at a school where one can do it for

you – or you can pre record some of the frequency changes and then listen to them again, not remembering the combination of frequencies of course.
Go through all the frequencies one by one, and listen to them for a few seconds each. Write down your names of the frequencies and now we can get started.

There are ten frequencies, and you have just heard them all. Now choose a random one and increase it by 12 dB. And then back to no increase again. (hint: write down the frequencies which you increase)
Write down on your paper, which frequency was it that you just heard increased.

Go to another frequency and do the same, until you have been through ten frequencies. Statistically you should have heard each of the frequencies once, but in practical you probably heard the same one twice or even three times.

Compare your own heard frequencies with the ones that was actually played.
It's ok not to get it all 100% right, the key is to understand what you actually heard, and being off by just one is basically fine.
After, even only a few attempts, or maybe even only one or two, you should be able to get 7 or 8 out of 10 right.

Now, doing this only once doesn't help you much. But doing this, which takes only 5 minutes, each morning or at the start of class – will greatly increase your listening skills and spotting that EQ frequency super fast.

Excersise one – expanded:
Now just as the "non expanded" version of this exersise, try lowering the increase in dB to only 6 dB. It should be much harder now.

To make it even harder, instead of adding to a frequency, do the exersise again, but this time choose a frequency and remove 12 or 6 dB.
Listening to an increase in a frequency is, compared to removal of one, easy! And if you can get it right on which frequency was actually removed then you are getting very good.

Now – instead of listening to pink noise and increase and decrease all the time, try putting on a piece of music. Now as much as you may like noise artists, electronic music etc. Then try to put on something cheesy pop music like instead. Mainly because this type of music contains all the frequencies you need, as well as focus on voice, good solid bass production etc.
Now increase or decrease, by 6 or 12 dB the same frequencies in the music track and see if you can still spot them.

The reason for this, is that music and sounds is what you will be using this tool for.

Basically the entire tool is your ear, but since this is about training your ear – automatically your skills within making music and sounds will increase.

Interesting Reading Material

Andy Farnell – Designing Sound. MIT Press.
Curtis Roads – The Computer Music Tutorial. MIT Press.
Curtis Roads – Microsound. MIT Press.
F. Alton Everest – The Master Handbook of Acoustics.
E. Bøgh Brixen & Voetmann – Praktisk elektroakustik. Forlaget Fog.
The Soundscape – R. Murray Schafer. Destiny Books.
Music, cognition, and computerized sound. Edited by Perry R. Cook. MIT press.
The Ultimate History of Video Games – Steven L. Kent. Three Rivers Press.
Sound Design – David Sonnenschein. Michael Wiese Productions.

References
Basically everything here is written straight off my head, but.
All the people who taught me something usefull and all books that I truly learned something from reading and studying, are my references.

Mark Grimshaw, University of Aalborg. Game sound theorist and audio ecologist. Had some lectures at the IT University of Copenhagen, Denmark.

Gordon Calleja, IT University of Copenhagen, taught me everything i know on the subject of immersion and a completely different perspective and approach on games and their theories.

Wayne Ziegel, Taught electronic music composition and did a ton of seminars at the Academy of Music / Jysk Musikkonservatorium, Aarhus Denmark.

Andrew McKenzie, my electronic music composition teacher for two years, and did a ton of psychotherapeutic seminars at the Academy of Music / Jysk Musikkonservatorium, Aarhus Denmark.

Bjørn Christiansen / Bjørn Svin, my electronic music composition teacher my last two years of study at the Academy of Music / Jysk Musikkonservatorium, Aarhus Denmark.

Henrik Winther – My acoustics, audio engineering and electronics teacher at Academy of Music / Jysk Musikkonservatorium, Aarhus Denmark.

Aleks Dubinski – A good close friend, audio programmer at IO Interactive, later at Unity, gave me some awesome lectures on C# coding and tons of tool design ideas.

Jakob Schmid. Ma.Computer Science – Audio programmer at Playdead. A good and close friend, taught me a lot of C# and we worked on fmod and Wwise together.

Morten Breinbjerg, My professor at the University of Aarhus, teaching audio, interactive audio, max/MSP and a lot of other things.

Curtis Roads, who's book The Computer Music tutorial, has been a great learning experience for all my years of study.

Henrik Munch, Sorten Muld, my electronic music composition teacher for a few semesters, as well as teaching me a lot of electronics, synthesis and more.

Nick Collins. University of Chicago, for doing a fantastic workshop on electronic music and electronics, components and building of such at the Academy of Music / Jysk Musikkonservatorium, Aarhus Denmark.

Derek Holzer, giving awesome workshops on soundscapes, electronics, sound ecology and much more at the Academy of Music / Jysk Musikkonservatorium, Aarhus Denmark and University of Aarhus.

Thomas Knak, System F3, for being a great inspirational source and giving a great lecture together with me at the Music School of Engelsholm. On sampling, sequencing and sound design.