As listening human beings we are constantly surrounded by "clouds", diffuse acoustic fields. Their composition determines, how voices and music appeal to us.
Jürgen Strauss on sound images like sugar and those that leave us cold.

Enveloped in sound

Jürgen Strauss partly works "in the clouds". At ETH Zurich, he is involved in planning an immersive lab, that will enable the acoustic properties of a room to be tested even if that room does not exist. “Auralisation” is the term used to describe the artificial audibility of an acoustic situation. Not only for future buildings, but also for the acoustics of buildings that no longer exist, important reference points can be gained in this way.
After training as a physics laboratory technician, Strauss specialised in the development of sound reinforcement systems for studios, concert halls, churches, museums and cinemas. He combines these electroacoustic fields of activity with room acoustics. At the ETH, he reads on the topic of "Historical and Systematic Foundations of Acoustic Architectural Design" and teaches his students about Ovid, the Epic of Gilgamesh and Plato.

Which phenomenon in acoustics can be described by the term "cloud"?

An acoustic condition, in which someone speaks or plays music and someone hears, is characterised by the fact that in a temporal sequence, one first perceives the direct sound. If the event takes place in a room, then with a temporal offset - depending on how the room is constructed - the floor reflection follows first, then the side reflection, third the ceiling reflection, fourth reflections from all directions, which overlap. This is called the acoustic room response. Every listening position or every listening situation is characterised by a very specific relationship between direct sound from the source and acoustic room response. Due to the superposition, the reflections form what we technically call a diffuse field. This diffuse field can be described as a cloud.

What we technically call diffuse field can be described as cloud.

In layman's terms, wouldn't "cloudless" hearing be much more desirable?

In our living rooms, at our workplaces, on the train, we are used to the space providing a specific response. When the world does not respond, many people feel as if they are suspended in the cosmos. This can very quickly become uncanny. One loses orientation. People who work in low-reflection spaces like this, in sound studios for example, soon notice that they have lost their sense of time. Because normally you move in time through a sequence of rooms, each with its own spatial response. The second sensation lost is the sense of space.

So the "cloud" anchors us in space and time?

Exactly, our perception, and in this case the auditory perception process, constitutes both the sensation of space and time through reflection! This becomes immediately tangible when we visualise how blind people orient themselves. There we see, that they create amazing localisation performances and spatial impressions by evaluating the reflections in space. So apt, that they can move around tables, recognise steps. They make a snapping sound, giving them a rather three-dimensional impression, right down to the size of teacups. We can all do that too, it's a matter of practice, but we just don't, because we orient ourselves visually. Nevertheless, we also find our orientation by evaluating reflections and have an impression of space. We always know whether we are in a small room or a long corridor, etc.

Are there further parallels to visual perception?

The lower the direct sound is in relation to the diffuse field, the more one has the impression that the sound is soft, decontoured, that gloss effects are dampened. This is a complete analogy to light and becomes clear when you compare camera lenses, for example portrait cameras from the fifties. The direct light that comes in and reaches the photographic paper directly at the back is, in a sense, our direct sound. And the light that is scattered in the lens and reflected at the edges of the lens creates diffuse light. The Leica company had a lens with a very specific grind, that was extremely well suited for portrait photography, because it had a slight attenuation of all contours, colours ...
... wrinkles ... . The opposite was also used for portraits, for example by Helmut Newton with the naked combat heroines he photographed, and that - this is part of this image aesthetic - with Nikon. These lenses try to capture pure direct light. That's why you can see the last pore in these pictures, whereas with the Leica lens the image is soft and slightly blurred.

Are there also such opposing influences when it comes to recording music?
At the time when Leica lenses were very popular, stereophony was also emerging, and it really took off in 1954. The English label Decca made its first stereo recordings in Geneva at the Victoria Hall. They are characterised by good direct sound, a beautiful clarity, you can hear the contours of the sound images well, but you also have, quite intensively, an acoustic spatial

response: Therefore, you can hear the unity of the performance space and the orchestra.
In further development, however, symphonic music in particular is not only recorded from a certain distance, but at the same time the microphone is moved much closer to the musicians. So-called support microphones are used, which are placed in the orchestra, half a metre or a metre away from the instrument. This is the so-called multi-microphone recording technique. The EMI label and later especially Deutsche Grammophon are known for this.

In principle, this creates a "group picture" of the orchestra from a distance of 5, 6, 7 metres, including a part of the room reflections. All the signals that are drawn directly from the orchestra with the support microphones are mixed into this image. This creates a sound image that is actually completely unenjoyable. Translated into the concert situation, this would mean: With one ear we sit on the musicians' laps, with the other in the back of the hall. In order to homogenise this, a sauce is poured over the mixture, so to speak, with an artificial reverberation that does not come from this room at all. And only in this way is the sound image created that we normally hear on recordings. This is what made the Berlin Philharmonic under Karajan so famous. You hear precision, detail. The timbre of the instruments is super clear and super distinct. But this is a sound image that you will never be able to hear in any concert hall on the planet.

Do we really like something so unnatural?

The great wealth of detail that gives an impression of proximity, obviously has a great appeal. This has been observed since the introduction of sound recordings. When you see illustrations of Caruso singing into a horn. He is perhaps 30 cm away from it and the accompanists are also as close as possible. This funnel is, indeed, the microphone, which in this case records an almost pure direct sound.

So Caruso sings right into my ear?

Exactly! He is very close. That creates a form of intimacy. When it comes to language, you can follow effortlessly. This possibility of sitting in the front row, or on the stage, has been very vividly embraced by the audience from the very beginning. Maybe, because we are such experts in regards to voices from birth and even before that. It's like sugar, we can't get enough of it.
And to finish this beautiful context off with an ugly example: Adolf Hitler. He had a microphone called "die Hitlerflasche", made by Neumann in Berlin. You can see it in many photographs and film footage: a cylindrical bottle with a round top. It almost had the quality of today's microphones. And then there was “der Volksempfänger” (public receiver), which was also an innovation. There are photographs that show how it was used. Sunday afternoon, 4 p.m.: The Führer is speaking. The speaking distance to the microphone is perhaps 30 cm, the transmission via radio can almost be neglected, and the Volksempfänger is on the coffee table with the German family, listening distance perhaps one metre, so that this Herr Hitler can be heard like another member of the family!

Even all pop music, "The Singer makes the Song", is always an almost pure direct sound, interception distance a few centimetres. This may later be faded or spatialised, but in principle the singer is at the centre of the action.

And what if we don't listen to recordings, but go to a concert?

An audience, that is used to this kind of sound aesthetic, because they listen to pop, radio, recordings that are produced in this way, also goes to a concert hall with this expectation. And then they are always slightly disappointed at first, because the music is not as clear, not as distinct, not as brilliant, you don't hear as many contours, not as many details. After a while you get used to it, but you have to swallow the bitter pill. Because at the medium listening distance in the Tonhalle Zurich, the Stadtcasino Basel, the Casino Bern, the KKL, the diffuse field dominates over the direct sound. The cloud effect dominates the event, whereas in all types of recordings the direct sound dominates.

So the Karajan method is still common today?

That has actually become the standard. There are other recordings where a so-called main microphone is used to record an overall picture from a certain distance. It sounds very different, it's very appealing, but the number of sales clearly show that it's much less appreciated.
So we are all very well practised in listening to such recordings. And when we go to a pop concert, a symphony concert or the opera afterwards, the organisers and the architects who build such spaces are well advised to try and provide sound the way that we are used to.

So the recording does not follow the concert impression, but vice versa?

That is a retroactive effect, that can be proven very well. As early as the 1930s, the technical level of electroacoustics had an effect on room acoustics. A famous example is the Elliot Hall of Music at Purdue University in India. This is a hall, that was first designed for symphonic music and tried to achieve the clarity and distinctness that could be heard on the mono recordings of the time. In principle, the shape of the funnel, like this Caruso funnel, was adopted as the floor plan. The orchestra was positioned in the narrowed side and the audience in the widening space. And that meant that the lateral reflections were perhaps reflected once at the opening walls, but very quickly redirected to the back. And at the back, a great absorption had been set up, so the reflections were swallowed up. In this hall, one could almost hear a pure direct sound.

Was the design successful?

The design worked. Clarity and distinctness were indeed very great, you could hear things you had never heard in concert before. But you also got tired of the hall very quickly. Because symphonic music lives very much from the interplay of direct sound and room response. And if there is nearly no room response, as in this hall, then it is unappealingly boring.

Did everything disintegrate into individual voices?

You could hear exactly how the violin bow flitted over the strings, you could hear every fine detail, not loudly, because it was almost only direct sound, the timbres were coherent, but it was absolutely dull because it did not involve the listener in any way! That only happens via the acoustic room response. That's why this hall had to be reworked. And so they installed loudspeakers that produced the missing reverberation, but secretly, not visibly, because that would have caused riots.

> Every conductor should be aware that the sound on the podium is completely different, than what you hear at the back of the hall.

Why is that?

It does not fit with our idea in any way that speakers are used for the best of sound for singing and the best of orchestral sound. We really want to hear the pure nature of the sound, of the expensive wood. We don't want to not be affected by technology. And that's how we felt even back then.

And technically that was already possible at the time?

Yes, one added in an electro-acoustic way, what was missing in a room-acoustic way. That can work very well, as in the Tonhalle Maag. Today you can expect an audience to do that, at least temporarily. But if you wanted to install a few more loudspeakers in the old Tonhalle hall to achieve even better acoustics, that would not be possible. It is completely unthinkable for the opera. It's about individual voices, about exactly this timbre, that's what we spend a lot of money on - if something comes out of a loudspeaker, the experience is ruined!

Have the acoustic characteristics of concert halls also influenced compositions?

I have been pursuing a research programme for a few years now. For that I have been looking for a case in which a composer composes for a particular room, because he knows the acoustic response of that room. This applies quite precisely to the situation of the young Joseph Haydn, who in 1761 was hired as director of music by Princes Esterházy. He finds a baroque banqueting and dining hall that had not been built as a concert hall at all. He notices that this hall has a very lively acoustic response, i.e. it reverberates quite strongly, with a reverberation time of around 2 seconds, and it generates an even longer reverberation, around 3 seconds, with low tones. It is relatively narrow, around 12 metres, quite long, around 30 metres, and quite high, around 10 to 12 metres. And all the music is really only ever directed at one listener, namely the prince. In his place, it is supposed to be exactly what the composer has intended to be heard. My research is to prove that Haydn considered the acoustic room response in his compositions very early on, especially in the period from 1761 to 1768. In order to make certain tonal phenomena audible at all. If this connection can be shown, then it would be proven that there is a direct connection between the genre of the symphony, which Haydn founded, and the so-called shoebox form of concert halls, long, narrow, high.

My research is to prove that Haydn considered the acoustic room response in his compositions very early on.

What about halls like the Elbphilharmonie?

They simply don't work for symphonic music! The model for the Elbphilharmonie is the Philharmonie in Berlin by Hans Scharoun with the central arrangement of the orchestra. The audience is therefore placed around the orchestra. In this context, there is a statement by Hector Berlioz, who was an exceptionally sensitive arranger, interpreter and - to put it this way - music producer: an exceptionally listening musician. He said that he understood the orchestra as a "hearth of sound". A fire gives off light and heat in all directions, so why not sit around the orchestra in all directions? Berlioz also organised mass concerts. If you want to supply as many people as possible from one orchestra source, then the round shape makes sense, because then the average listening distance is as short as possible. But Berlioz then describes very vividly what happens: If you are very close and the front is played with inspiration, then it is impressive, you are gripped and the affects rush through your body. But if you are far away, the music is quieter, it leaves you cold. You are far away from the fire, the passions are not awakened.

What's the point of an Eroica at room volume? It needs to thunder.

In fact, the music is louder on average and you can see the orchestra better if you sit all around. This is also relevant for the concert impression. In the shoebox-shaped concert hall, however, there is the so-called source widening effect: shortly after the direct sound come the lateral reflections. These early and strong reflections from the side seem as if musicians were there as well. This not only makes the signal louder, it also makes it seem wider. In the Elbphilharmonie and Scharoun-Philharmonie, these lateral reflections do not exist at all, or they are very weak. And what is the point of an Eroica at room volume? It has to thunder.

Hector Berlioz said that he understood the orchestra as a "hearth of sound".

Why was it built like that anyway?

Today, when you build a concert hall, you want to accommodate as many people as possible. That was the case with the Philharmonie de Paris, the Auditorium Parco della Musica in Rome and also with the KKL. These halls can hold around 2000 people. But if you want to create the same dynamics for 2000 people as Haydn did in his hall for 40, you would have to bring in an orchestra of several hundred musicians. But that can no longer be coordinated. The halls are simply getting too big. The acoustic room response no longer contributes much to the impression of loudness; the reflections come too late, then they are already too quiet.

In the KKL, however, this is a much less dramatic problem, because this hall is really exceptional in one respect: when all the doors are closed, it is enormously quiet, there is practically no sound in there that would be generated by passing ships, trains, cars. Usually, especially for low frequencies, the coupling is via the ground, that's where the structure-borne sound comes in, reaches

the walls and makes them vibrate. They are then like a membrane and generate sound in the room. The KKL is constructed in such a way, that it is almost completely decoupled from its environment. This makes it possible to play pianopianissimo and still hear it, because there is no noise to obscure it. Here, the cloud could be used again as an image: Such a basic sound is, to a certain extent, a cloud that pushes itself between us and the sound event and partially or completely obscures it. In this sense, the KKL is a cloud-free zone!

Musicians rarely address such contexts. Why is that?

I would call it a loss of know-how or sensitisation. We know quite well that it used to be common practice to adapt the performance to the acoustic conditions of the room. To a certain extent, this is done as a natural thing today, just as when we speak in very reverberant conditions, we articulate better, stand close together, reduce the dynamics of the voice and speak rather slowly. But this hardly ever takes place in a reflected way. Conductors should be aware that the sound they hear on the podium is completely different from what they hear back in the hall. There is a great distance in between, and that is where our diffuse field, our cloud, comes into existence, which has an enormous influence on the aesthetics of sound.
The direct impact of sound, which the musicians have experienced, is also essential. They have the feeling that if the music is clear and distinct, if they can hear themselves and the other players well, then it is good. Yet this has little to do with what the audience hears. The training courses do not necessarily sensitise them to such questions. Therefore, it can be said that paradoxically, especially in symphonic music, this pièce de résistance of musical room acoustics, there are often people who are satisfied that their playing is in accordance with the score and that the tonal balance and dynamics are satisfactory in relation to the conductor's podium.

Surrounded by clouds

Summary: Jean-Damien Humair - At the Department of Architecture at the Zurich University of Applied Sciences, Jürgen Strauss is planning to set up an immersive laboratory to test the acoustic properties of a room, even if it does not exist. This process, called auralisation, makes any acoustic situation audible. This will make it possible to assess the acoustics of future buildings, but also of buildings that no longer exist.
Jürgen Strauss explains that when you listen to someone talking or playing music in a room, you first hear, in chronological order, the direct sound. Then you hear the floor reflection, the side reflections, the ceiling reflection, and finally the reflections from all the overlapping directions. This is called the room's acoustic response. All these reflections, called reverberation, can be described as a cloud.
This reverberation is present everywhere, in our living rooms, in the office, on the train. A place that produces no sound reflections is distressing very quickly. We lose our bearings. People who work in such places, such as recording studios, quickly lose track of time - because we are used to moving around in different places with different reverberations over time. The second phenomenon is that we lose the sense of space. The sound cloud thus anchors us in space and time. Blind people orient themselves in space by evaluating sound reflections. They are able to identify a table or a staircase, up to the size of a cup of tea. We are less developed than they are, but we can also always tell by ear whether we are in a small room, a corridor, etc.
In 1954, Decca made the first stereophonic recordings at the Victoria Hall in Geneva. They were characterised by good, direct sound, good clarity, but you could also hear the reverberation of the hall quite intensely. Later on, proximity microphones were added, placed at about 50 centimetres from the instruments, to complement the main microphones located at 5, 6 or 7 metres from the orchestra. The EMI label and later especially Deutsche Grammophon specialised in this type of sound. This gives an impossible sound image, as if you were sitting on the musicians' laps with one ear and staying in the back with the other. To make it consumable, an artificial reverberation is added, which has nothing to do with the hall.
The Berlin Philharmonic became famous for this under Karajan. In his recordings you can clearly hear the precision, the detail of each timbre, but it is a sound image that you will never hear in any concert hall on the

planet. Yet our ears love this kind of sound, this closeness. This is the case with today's pop singers, who are recorded at a distance of a few centimetres and thus literally sing "into our ear".

This was also the case with Adolf Hitler, who gave his speeches 30 centimetres away from the famous Neumann microphone and who, through the radio, spoke to every German who listened to him.

Today, some listeners are disappointed when they go to a concert, because they do not get the sound aesthetics they are used to. The sound is not as clear, as brilliant as on the recordings. So from the 1930s onwards, the acoustics of the halls were adapted, so that they produced less reverberation. This was the case, for example, in the famous Elliot Hall of Music at Purdue University in Indianapolis, which opened in 1940. You could hear the bow of the violin rubbing the strings, distinguishing each timbre of the orchestra, but it was absolutely boring because it didn't involve the listeners. So we added loudspeakers that produce a false reverberation, but we hid them, because otherwise it would have caused a riot: the audience wants to hear the natural sound, the precious wood of the instruments. Only in temporary situations, such as the Maag Halle in Zurich, one tolerates fake reverberation.

In his work, Jürgen Strauss also tries to show that some composers have adapted their writing to the acoustics of the hall in which their work was to be performed. This could be the case, for example, with Haydn, who wrote for the Esterházy princes in their dining room, the acoustics of which were not at all like those of a concert hall of that time. If this relationship could be demonstrated, it would be the evidence of a direct link between the symphonic genre, which Haydn created, and the so-called shoebox shape of the concert halls built later: long, narrow and high, like the prince's dining room.

But today's concert halls are designed for 2,000 people, not for 40 as at the Esterházy court. To have the same volume of sound, you would need much larger orchestras. However, some venues manage to cope with this, such as the KKL in Lucerne, which is almost completely isolated from the environment. No extraneous sound reaches it. You can play pianissimo and all the listeners can hear it. In a way, the KKL is a cloudless zone.

Musicians are rarely aware of this issue. We know that in the past it was usual to adapt one's playing to the acoustics of the venue. To a certain extent, this is still the case today: we articulate better when we speak in very reverberant conditions, and we stand closer together. But this is hardly reflected in the musicians. The conductor must realise that what he hears from his podium is not at all the sound that the listeners in the hall perceive. The musicians have the impression that the sound is good when they can hear their musicians, but here too they have to realise that this has only little to do with what the audience can hear.