TIMING, LATENCY AND LIVE PERFORMANCE

Robert Klimek
Humanities, Arts and Social Sciences
Music Program
Colorado School of Mines
Golden, Colorado 80401
rklimek@mines.edu

Catherine Skokan
Humanities, Arts and Social Sciences
Music Program
Colorado School of Mines
Golden, Colorado 80401
cskokan@mines.edu

Abstract—How important is synchronous timing in a performance? In the world of science, engineering and the arts, we cannot avoid it. In the arts, it is a living and vital concept between composer, performer and audience. During the Pandemic, with an increase in virtual meetings and performances, made us more aware of timing issues and introduced us to a path with many questions about latency and affect between audience and performer. The hand waving, head nodding and eye gestures of Renaissance musicians led eventually to the classic conductor’s baton. These all helped ensembles of dissimilar instruments and/or voices, conquer problems of ‘sounding’ simultaneously, in order to create accurate harmony. Music performers, of all styles and types, consistently alter their attack times to create a synchronous performance for not all instruments/voices immediately sound out a fully developed tone at the same time. Recording techniques (one solution) can ‘doctor’ this problem, as sound engineer realigns tracks, and starting points, when necessary. But are there other new latency solutions available to us today and discovered during the pandemic year and a half?

Keywords—timing, latency, delay in sound

I. INTRODUCTION

In January 2020 MINES worked on the acoustical and digital challenges of timing through the shared experience of the production of 2 live 90 minute shows with the 2020 National Western Stock Show in Denver, Colorado. The task was to produce a seamlessly timed performance for a 20,000-person audience, using singers, concert instruments, a jazz band, horses, adult and children riders, and actors. Our task was to produce a seamless, “real-time”, ebb and flow experience for the audience despite issues caused by curious young horses, to riders, and actors all having separate tempi, the task was to produce a seamless, ‘real time,’ ebb and flow experience for the audience.

COVID has made all performers hyper-conscious of timing and latency issues. This issue has always been a part of live, real time performances; however, quarantines, separation and isolation have spotlighted it. Now, instead of being something everyone does naturally, it must be factored into each rehearsal and performance. Without the push of COVID, Zoom might have been a handy program to use from time to time. But now it has become an office, a party room, a family dinner, and our local brewery. Zoom taught us that sending signals back and forth over the internet took the analog precision of performance and timing away. Prior to Zoom, musicians needed to make sure their instruments were in tune, etc. With Zoom and COVID, one had to ‘align’ with the another. And the other, might have poor WIFI, a low-cost microphone and small or only computer-monitor speakers.

Can compositional or performance artistry overcome this problem? Can sound and electrical engineering eventually overcome this problem? What’s been tried? What was seemed successful, but proved to be more costly? Will the return of live performances make everyone step back to prior times? Or, has something been learned during COVID which will drastically change how we run live and internet performance? What’s next?

II. CSM INTEREST IN LIVE PERFORMANCE ISSUES

Synchronous timing of performances is an engineering issue as well as an artistic issue. The Colorado School of Mines (Mines) is a public research university devoted to engineering and the applied sciences. Our degree offerings are all technical. However, approximately 15% of the students also participate in the performing arts: Band, Orchestra, Choir, Jazz Band, Theater, and small ensembles. This unique combination has taken an interest in the topic of timing, especially during 2020 and COVID challenges.

At Mines [1], our degree offerings include the traditional earth sciences which serve as our historic foundation, as well as engineering and science/math arenas. These include geologic, geophysical, mining, and petroleum engineering as well as materials science (metallurgy), computer science, mathematics, and chemistry. Degrees are also offered in the traditional engineering topics of civil, mechanical, electrical, and chemical engineering, to name a few. Our students must complete a humanities component in their engineering curriculum, and many chose the performing arts to partially fulfill this requirement. The performing arts fall under the non-technical department umbrella called Humanities, Arts, and Social Sciences (HASS). This department offers courses in language, international affairs, ethics, music and performing arts, literature, and philosophy. In addition to the performance opportunities, Mines also offers individual music instruction, classes in music theory, composition, and history Our students can receive a minor in Music Technology or integrate a focus in music technology as part of a general B.S. in Engineering degree.
As a highly selective school, one in ten applicants enroll at Mines. The student body includes about 4900 undergraduate and 1500 graduate students. Almost 30% are female and 11% are international. Minorities count for 19% of the student population. Our performing arts classes draw from the entire student population and often attract higher percentage of females than the school average. Our music classes are popular, usually fully enrolled, and often have a waiting list. During non-COVID times, the class sizes range from 35-45. Our band has 115 students, orchestra has 78 students, and choir with 75 students. The theater class regularly enrolled 30 students and had a waiting list. During COVID, the class sizes have been smaller because many students chose to remain remote for engineering classes.

III. HISTORY OF CAUSING AND SOLVING ‘LATENCY’ PROBLEMS

Early, small ensemble, instrumental music began and ended through the simple use of eye contact or casual gestures. One member simply volunteers for this role. As Baroque music became more complicated and ensembles grew larger, it soon required a musician whose sole purpose was to keep the tempo of the music consistent — a conductor. One of the earliest was the Baroque composer Jean-Baptiste Lully (1632–1687). He would stand before his orchestra with a heavy, spiked staff, beating time by pounding the staff on the floor of the conductor box.

The heavy staff proved to have two disadvantages. It was annoying to listeners and ultimately ended in Lully’s death. A distracted Lully accidentally pounded his foot, which eventually turned gangrenous. The spiked staff was quickly replaced with a rolled-up score, much softer and less dangerous, and finally a much small version of Lully’s staff, a baton. Something unique in ‘real time’ performance began to take place.

The baton allowed for more graceful gestures, and thus the conductor took on a more central role in music performance. The conductor could shape the music, adding another ingredient to performance — interpretation. Most orchestras play ‘behind’ the conductor. In short, they wait a little for the conductor indications before they sound their instruments. So, the conductor no longer ‘beats’ the tempo to keep the instruments together. Another ‘real time’ understanding of live performance was created.

Consider this thought: large ensembles began to deal with the geography or placement of their membership on the stage. Sound travels to the listeners ear from the musicians closest to the front then the instruments in the back. The speed of sound in air is about 343 meters per second [2]. Therefore, sound travels about one meter in 3 milliseconds or in one foot per millisecond. The human ear can discern a lag over 20 to 30 milliseconds or a distance of 20 to 30 feet from sound source to ear. Major symphony orchestras have a space of about 35 feet by 35 feet or more. Orchestra musicians are aware this lag, as well as that of their instrument response. The attack, decay, sustain and release (ADSR) is different from instruments to instrument. For example, a violin’s attack is most immediate, a bell is even quicker. The tuba and bassoon have a lag time from the moment the air enters the instrument until a sound is produced. Musician’s, with the conductor, adjust, in real time, to compensate for this situation. Composer’s use groups of instruments to cover over this problem in their orchestra of works. Audiences accept the blending of differences in orchestration. Our brain compensates for slight differences in timing. Without these compromises and shifts, harmonies would be hard to achieve. Melodies would not hold together.

IV. A LIVE PERFORMANCE MODEL FOR HOW MUSIC IS CREATED AND RECEIVED

In the world of physics and engineering, we think of a pebble dropping in a still pond and creating expanding concentric circles according to Huygen’s principle. Now, think of a pebble dropping in a still pond and creating four concentric circles. The center circle, caused by the pebble striking the water is the Music itself, the created art before it is performed. The next circle is the performers. The next is the Audience. The final is Time and Space in which the performance occurred. As in the energy passing back and forth from the ripples cause by the initial pebble strike, each of the components back of a music performance have a similar energy that must pass back and forth to make the performance in ‘real time.’ “Music is not something “given” but . . . that which . . . rests on (the back and forth) agreement . . . (between) . . . composer, performer and listener.” In order for music to achieve a ‘live’ performance model it must be able to move and affect the performer to the listener and again.

The composer starts out with a desire to be heard, with a desire to move a performer to eventually perform their concept so that their music ‘moves’ a people to experiencing an expected response. Affective music performance is akin to what happens when one drops a pebble into a still pond of water. The pebble creates a series of outgoing concentric circles. From middle moving outwards, these concentric circles can be labeled compositor, performer, audience and historical moment. From one circle to the next and back again, there is movement. As in the pebble drop image, ‘latency’ rules or ‘on-time rules’ that are agreed upon in between these concentric circles (composer, performer, audience and historical moment of performance). The performers and listeners do not, “discuss them, . . . rather ... they have absorbed them. “And by continual group practice they agree to them.”

This agreement of playing together, in time and starting and stopping as one, is also something the audience agrees on hearing. The audience comes to the performance knowing they will hear something new, but this new is within the agreed upon structure of the performers. Will this be a parade with a marching band? Will this be a jazz rendition in a smokey club basement? Will this be a choral and string piece in low ceiling church, or a gothic cathedral? All these historic places carry with them a memory and history which also empowers or affects back to the audience, the
performers and finally the composer. The ripple ‘affect’ you see, moves back and forth in ‘real’ time.

As music became amplified and played in large arenas, amplification came to be part of the concert presentation. Massive audiences wanted to get the most ‘bang’ for their buck, and performer’s amplification came to be a part of ‘rock’ concerts. Now a “disembodied voice, coming through a machine, . . . remains enigmatic to the audience.” “In time” music takes yet another step forward with the introduction of personal playback machines and ear buds. Now the listener has no ‘affective’ connect with the performer, nor does the listener give any affective feedback to the performer. The composer only gets a royalty check to let them know how well they are doing. The historical moment can be in a gym lifting weights or a walk about a park while reading text messages. Again, what makes the music real and ‘in time’ is no longer present. This was happening prior to COVID, which now has added another layer of complexity.

Finally, music becomes fully electronic (Electronic Dance Music EDM), and the light show become a most important source of ‘entertainment’. The performer moves dials. Someone else runs the lights. The audience acts upon agreed rules, but unlike a live performance of jazz improvisation, things can be easily and exactly repeatable. Some come because they already know the product. They like that. However, the first model we used of a community hearing music is far from this model.

So, the latency of choirs and organs and instruments in resonant renaissance cathedrals is solved partially by learning to play and sing without your ears, and much of the ‘latency’ or real time performance affect, is accepted by the audience, in a large part due to the historic place and moment of the performance. It was agreeed on by performers and audience how this would sound and take place. Musician and audiences, however, could still in a sense ‘ripple’ back and forth and support each other to make the music present and in real time.

Network Produced Music (NPM) posed and even more unique problem. There are no ripples, and everything now depends on your computer, your computer speed, your headset or speakers, your microphone; your interface; how far away you are from each other electronically.

In the recording or transmission of sound, latency can become a difficult issue. This has become especially evident as we have relied on internet communication platforms such as Zoom during COVID. Our sound is produced and picked up by a microphone. We can get the microphone close to the sound source to minimize delay. The sound is then converted from an analog signal to a digital signal. Software engineers have been successful in minimizing this delay. Then the signal is sent through a transmission medium such as the internet. The signal is then converted back to an analog signal and through a speaker to our ear. Latency is a function of the slowest of the units, in this case the transmission medium. A delay of over 100 or 200 ms or more is not uncommon. This delay is easily recognized by the listener.

In COVID, new questions about putting together live ensembles became important to consider. The only other option, besides ZOOM and similar programs was that of silence or nothing at all. In short, COVID was the death of live performance. No one wanted to go to that funeral.

V. TWO QUESTIONS STAND OUT

So, the problem with following the ‘in time’ live performance (pebble drop in quiet pond concentric circle affective performance model) lies in two areas and with two questions needing an answer. First, music prior to the COVID period strove to always fulfill the pebble/concentric circle model of live performance in one way or another. Post-Covid performance is already showing us that this model is changing rapidly, even when face-to-face performance is again possible. As an example, bands now can be made up of a single person recording in his/her bedroom, with no thought of live performance. An actual ‘live’ band doesn’t exist in this case. How will this affect how we understand live music performance in the future?

Next, as has been done in the past, do composers need to learn how to create a new type of music composition that considers the latency problem and works with it? Surely, composers can, and have during COVID, found a way. They, like the organists and choirs in large cathedrals, produce music that will bring performers and audiences into a ‘real time’ music performance moment. Already, composition students are tackling this problem, knowing the pre-COVID model is not dead, but no longer the only ‘affective’ model of ‘present’ ‘real time’ music performance. New music (non-pop forms) has always found it difficult to find an audience. This could now be an opening. A single performer could stand in your home and perform, while the rest of the ensemble joins them remotely, through the use of specific software and a mixing board. In this way, the pebble/concentric circle model is present.

Another question lays on the engineering side. How do we arrive at an acceptable latency to create a live ‘real time’ performance? Can engineers, through technology, overcome the latency problem of a lack of presence between performer and audience? What was tried? How successful were those tries? COVID changed many things. If those changes remain, how will this problem of timing be solved so the audience can feel the presence of a caring performer? We also might find out that the advanced world of technology cannot replace the sensory attuned performer artist. An endless supply of thoughts about the challenges of latency, working around it, and some possible working solutions, have arisen. Here are a few pertinent ones, from a very long list.

[3,4,5,6,7,8,9,10,11,12,13]

VI. CSM PARTICIPATES IN A ‘REAL TIME”
LEARNING PERFORMANCE EXAMPLE

All this thought brings us to the Denver, National
Western Stock Show 2020. The Colorado School of Mines Concert Band, Concert Choir and Jazz Band were invited to supply the music for their Night of the Dancing Horses show. In Denver, this would become one of the last live full audience performances for over a year and a half. It also gave our students an authentic experience in varying latencies in ‘real’ time. The object was to accompany a spoken narrative track, trotting horses of different types, live on horse performers, as well as ground performers from the ages of 6 to adulthood, and ethnic and classic dancers. This was performed in a sonically challenged arena with dirt floor, steel roof, an array of 12-inch speakers throughout the ceiling and a band split into three to fit into the arena. To add to the disorder of sound was a crowd passing through the upper part of the building to adjoining arenas. To this, the music was to bind everything together so that to the audience would perceive a single, choreographed, real time performance experience. Sound from the stage could only be heard from the ceiling speakers. The arena absorbed most everything from the stage. Even at fff, the Concert Band was barely heard 3 feet from the stage.

The challenge was to make a real, in-time performance that would follow the concentric circle model of the performers affecting the audience in an historical place/moment with the ripples from the historical place/moment and audience filtering back to the performer.

The solution, in a nutshell was in finding the tempo of all the performers, as well as the speaker system.

The different breeds of horse trotted to different tempos. Young horses were mesmerized by all the instruments on the stage and forgot the performance entirely at points. The instruments on stage all sounded at different times, however, the sound system justified all their differences as their performance occurred not in the room ambience but in the microphone to speaker arena. Tempos were varied to move certain groups along. At times the tempo supported the dancers. At other times, the horses were supported. The synchronous moments gave an illusion to the audience that all parts of the performance were synchronous. The narration was patterned, timed, and practiced filling in moments when the music and the horses could not coordinate. Lighting effects pinpointed visual downbeats and covered over non-synchronous moments.

This YouTube clip shows the finale from the National Western Stock Show. (beginning at 4:38 provides best example)

https://www.youtube.com/watch?v=XOoNn4lN3q0

VII. NO COMPLETE SOLUTIONS, BUT PROGRESS IS MADE

All in all, real time latency was achieved despite ‘real’ live variables. In this same way, through new compositions that ‘work out’ the problem of latency, as well as finding new pathways on the internet and through innovation in Apps, music performers are again looking at a problem of realizing that “…music is a full-body sensory experience making full use of sight, sound, touch, and collaboration to produce a real time, concentric circle model, experience.”[14]

REFERENCES