Spring 2015

The Rise of Reactive and Interactive Video Game Audio

Joey Pendergrass
California State University, Monterey Bay

California State University, Monterey Bay

The Rise of Reactive and Interactive Video Game Audio

Joey Pendergrass

MPA 475: Senior Capstone

Dr. Lanier Sammons

18 May 2015

Abstract

As video games have matured from a technical hobby to a fully-fledged art form, so too have the methods used to create them matured. In tandem with the constant evolution of graphical fidelity, gameplay mechanics, input devices, and storytelling techniques spanning the entire existence of video games, there has also been an evolution in the purpose, and implementation, of music and sound effects. Whereas early video games utilized simple melodies to attract players in arcades, modern video games have grown to utilize audio as a means for storytelling in itself. As the visual component continues to grow more realistic with each gaming generation, equally realistic audio has become increasingly important in creating a cohesive experience for the player. For this reason, many developers have turned to utilizing audio pipelines which allow for expanded mixing and audio processing techniques, resulting in the most immersive interactive audio experiences ever made.

<u>Introduction</u>

Throughout the course of the last four decades, video games have rapidly expanded into one of the most widely consumed media forms of all time. Outside of the most secluded of cultures, one would be hard-pressed to find a person who has not, at the very least, heard reference of the expanding presence of video games in our every day existence. With the evolution of video games resulting in the recent integration of gaming platforms into almost every form of consumer electronic device, many otherwise uninitiated individuals have gained awareness of video games as both a hobby and an art form. This is seen most notably in the technological advancements made to mobile telephones over the last fifteen years, with the “smartphone” currently being a major factor in the most steadily expanding gaming demographics. In fact, the Entertainment Software Association’s “2014 Essential Facts About the Computer and Video Game Industry” report states that 44% of video game playing individuals utilize their smartphone for the very purpose of gaming. Along with the growing embrace of mobile gaming, console- and PC-based games maintain a serious share of the total market, with each hardware iteration creating more immersive experiences through increased graphical fidelity, reactive and interactive audio, and the incredible storytelling that they facilitate.

In terms of the roles that music and sound design play within these gaming experiences, one can easily see its importance through analyzing the function of audio within other forms of media which also present the consumer with both visual and audible stimuli. As with film and television, which both rely on sound as an integral part of their overall consumer experience, the visuals of a video game desire to be accompanied by some form of soundscape, which, for the most part, does nothing less than completely enhance the immersion of the player. While it can be said, in some instances, that the accompanying audio is not necessary to the player's experience, the vast majority of released titles strive to prove otherwise. In my opinion, an impressive gaming experience relies just as heavily on the presentation of incredible audio and visuals, as it does on the game play mechanics and writing.

It is with this symbiotic relationship between the various components of a video game that the audio engineer finds their niche within the scope of its production. As the core skills necessary to create a good mix in any context are much the same, the task of mixing audio in the realm of video game development is not entirely different than that of mixing for other forms of multimedia, such as music, television, or film. The most notable difference between the work that an audio engineer performs in the creation of compelling video game audio and that of an engineer working in other industries lays in the non-linear nature of the video game experience. Whereas music and film audio are designed to be experienced in progressive, linear, time, with the same content being delivered to the consumer in the same order every time without deviation, video games are inherently fluid in the way they present content to the player, due to their reactive, and interactive, qualities. The audio must be able to react to decisions made, and actions performed, by the player, and for this reason, the audio engineer cannot approach creating and mixing modern video game audio in the same linear way that is used in other media forms. In order to accomplish the task of creating the extremely cohesive and immersive experiences that are made possible through the combination of gorgeous visuals and intricate sound design, audio engineers utilize various tools for the implementation and mixing of sounds which work in tandem with the rendering and physics engines, to create a system which interoperates on a level so efficient, the player can at times lose sight of the complexity of the game in which they are playing, as if it just exists organically, and is perceived in the same way as the aural and visual stimuli present in reality. These systems are impressively complex, and are being used to create many of the incredible gaming experiences that have been released in the last twenty years, as well as those which will engage gamers for years to come. However, to fully understand the influence that dynamic audio has had on the overall gameplay experiences of the modern era, you must understand the complex role that audio has held throughout the history of video games, as a whole.

The Role of Sound in the First Video Games

If one were to survey the entire breadth of released video game titles and platforms, it would be apparent that not every game has utilized audio in an interesting or compelling way, with some of the earliest "silent consoles" forgoing the inclusion of audio at all, such as the very first home video game console: the Magnavox Odyssey (1972), and other early "PONG" systems of the same era (Hatzithomas). However, in the same survey, one would also find that, while sound was not always a focal component in the development and design of early commercial video games, aural stimuli can be recognized as an important factor in video games from their earliest days in research labs throughout the 1940s and 1950s.

In order to accurately distinguish the first video game to utilize sound as a meaningful component, we must first be able to recognize the creation of the first video game in history, which is a topic shrouded in a bit of ambiguity, due, primarily, to various differing opinions regarding the qualities and components that are necessary within an example of electronic entertainment in order for it to fulfill the definition of a “video game.” Thus, to identify the earliest video games, there must be acceptance of a definition which adequately describes video games in a way that limits the question of which of the earliest predecessors to the easily recognizable video games of the late 1970s best fulfills that definition. In the interest of this survey, I have used the Oxford Dictionary entry for “Video Game,” which defines the term as, “a game played by electronically manipulating images produced by a computer program on a television screen or other display screen." In utilizing this definition to qualify the earliest video games, we can begin to trace the history of the art form back through the decades leading up to the first commercial gaming platforms, and effectively pinpoint the first game which utilized audio in a meaningful way.

One of the earliest examples of a computer program which entirely fulfills the Oxford definition for “video game,” *The Bouncing Ball Program*, developed by Charles W. Adams and John T. Gilmore Jr. for MIT’s Whirlwind I computer in 1949, may also be the very first game to not only utilize audio, but also to generate sound effects in reaction to its visual content. As Jon Peddie writes in his text, *The History of Visual Magic in Computers: How Beautiful Images are Made in CAD, 3D, VR and AR,* the program simulated a ball bouncing within a displayed box with a hole in the bottom plane, with the “player” adjusting the display’s controls in order to allow the ball to bounce in a way in which it would fall through the hole. Peddie states, “the computer was capable of primitive sound; using timed beeps generated by the console’s speaker, and would make a thunk like sound when the ball (a dot on the screen) bounced off things, or fell in the hole" (82). The argument can be made that *The Bouncing Ball Program* existed primarily as an example of computer based physics simulation, however, it fulfills the definition of a video game set forth by the Oxford Dictionary, regardless of its purpose. Moreover, the user of the program was fully responsible for modifying the parameters within the program which determined the ball's trajectory, and as a result, were granted with a sense of victory when the ball's path resulted in its passage through the hole in the bottom plane. Through this mechanic, *The Bouncing Ball Program* can easily be recognized as game-like, in the sense that the user could utilize the program for challenge based recreation and competition. And with the Whirlwind I computer being capable of generating tones, which the program coincided with the ball's movements on screen, it also certainly qualifies as a game with reactive audio content. Although the sounds in *The Bouncing Ball Program* were extremely primitive, they set clear precedent for reactive game audio, which would be seen in many of the commercial video games of the next few decades.

As I mentioned above, well before the silent home consoles of the early 1970s were released, video games existed within research facilities and technology exhibitions around the world as experimentation in logical computing. Many of these games were designed to demonstrate the capabilities of emerging analog and digital computational technologies of the time, often serving as examples of artificial intelligence, as witnessed in Breyer Bettencourt's *Bertie the Brain*, a four meter tall machine built in 1950 which was capable of playing Tic-Tac-Toe against a human opponent with near unbeatable skill, and the Nimrod computer, built in 1951, whose purpose was to introduce people to the principles of digital computers through direct interaction; in the form of competing against the computer in the ancient mathematical strategy game known as *Nim* (Bateman; ”NIMROD"). Both of these devices fall well within the bounds of the Oxford Dictionary definition of a video game due to their use of dynamic visual feedback during operation, and are therefore held in the record as two of the earliest video games to ever be created. However, they are also both ultimately little more than intelligent computer systems designed to emulate the human mind in the contest of previously existing skill based tabletop games. For this reason, they tend to be discredited as true video games, with a classification such as “Artificially Intelligent Computer-Based Competitor” seemingly a better fit. Yet, in the interest of the definition, they are entirely viable.

While *Bertie the Brain* and the Nimrod computer are both milestones in the development of video games into the modern era in their own right, they did little to influence the implementation of audio as an important component to game design, as they were both silent. All the same, silence was not a common trait amongst the entirety of their contemporaries; though, in some cases, the presence of reactive sound content was not an intentionally implemented feature of design. This can be witnessed most apparently in another commonly regarded contender for the title of the very first video game, given the assumption that the systems described above do not fulfill a more refined definition for video games than the one I have supplied from the Oxford Dictionary: the aptly titled, *Tennis for Two.*

Developed in 1958 by William Higinbotham at the Brookhaven National Laboratory in Upton, New York, *Tennis for Two* holds an important distinction in the era of early video games, in that its creation was the result of Higinbotham's simple desire to invent a source of electronic entertainment for guests who would visit the Brookhaven Lab. According to an article on the Brookhaven National Laboratory's website regarding the creation of *Tennis for Two*, Higinbotham once wrote, "it might liven up the place to have a game that people could play, and which would convey the message that our scientific endeavors have relevance for society" ("The First Video Game?"). It is important to recognize the purpose for which *Tennis for Two* was invented, as unlike *The Bouncing Ball Program*, *Bernie the Brain*, and the Nimrod computer, which were all built to demonstrate developing processing and artificial intelligence capabilities, *Tennis for Two* was designed specifically for the sake of providing a game to be played amongst people. For this reason, many have considered *Tennis for Two* to be the first genuine video game.

Assuming the mindset that *Tennis for Two* is most deserving of the title of the first video game, the argument can be made that it is also the first video game to utilize sound as a gameplay component, albeit completely as an uncontrollable result of its analog design. *Tennis for Two* consisted of an analog computer built up of capacitors, resistors, and mechanical relays, which utilized the cathode-ray tube of an oscilloscope to display a two dimensional side-view of a tennis ball court, with a bright, moving, dot representing the tennis ball. As players utilized their individual controller, equipped with a single rotating dial and button to adjust the angle of the ball and trigger both serves and volleys, the analog computer would output changes in voltage to the oscilloscope, which would result in the movement of the ball on screen in relation to the parameters put in place through the player's interaction with the controller ("The First Video

Game?"). Through the adjustment of the ball's trajectory, players were ultimately able to participate in a fairly realistic electronic interpretation of a standard singles tennis match. The analog computer utilized for *Tennis for Two* was simply programed to adjust the voltage of the output which was connected to the oscilloscope, in order to display related ball movement on screen, both in reaction to player dial positions, and also through a complex system which determined when the ball would strike either the ground or the net, changing its trajectory accordingly.

The computer did not, however, feature an intentionally implemented system to provide audible reinforcement to gameplay. Yet, due to the nature of the mechanical relays used in the computer, which reacted to incoming electrical information to make physical changes in the routing of current through the circuit in order to generate the output voltages needed to visually render the arc of the ball, either from a hit at the angle given by the player's dial position, or a resulting change in trajectory due to the ball hitting the surface or net, the game featured a simple sound component. As a result of their switching, the relays would produce a fairly loud percussive sound, which in turn, exhibited sonic qualities similar enough to the sound a tennis ball would make when being struck, or striking a surface, that the sound was easily recognizable as holding correlation to the actions of the ball on the oscilloscope display. The sound of the relays switching the circuit can be readily heard in Brookhaven National Laboratory's video, "First Video Game," which accompanies the *Tennis for Two* article. In this short video, which outlines the re-construction of the *Tennis for Two* computer for Brookhaven's 50th anniversary in 1997, footage is shown of the re-creation being utilized for game play. Although this sound was purely a result of the analog components utilized in *Tennis for Two*'s computer, the video clearly exemplifies that the correlation between the ball's movement on screen and the switching of the relays qualifies the resulting sound as a reactive sound effect. Although it was not purposeful design that granted *Tennis for Two* the esteem of being one the first video games to utilize sound as a gameplay component, I firmly believe that the audibility of the mechanical relays served to subtly enhance the gameplay in a way that only sound can; an enhancement that would continue to develop into an integral part of video game design into the modern age.

Audibility in the Arcade Era

In the decades following the development of such incredibly innovative and archetypal video games as *The Bouncing Ball Program* and *Tennis for Two*, the interactive experiences aligned with video games as a whole began to escape the small audience of science and technology minded individuals within research laboratories and expositions, and enter the consumer's world as a groundbreaking source of entertainment for all to enjoy. The 1960s ushered in the development of video games for the singular purpose of providing meaningful diversion, with titles such as *Spacewar!*, developed for the PDP-1 computer by three MIT students in 1962 as a means to represent the full extent of processing power available with the machine, serving as the inspirational bridge between the lab and the arcade. While it ultimately existed as a product of research, proving insignificant in the breadth of video game audio due to its silence, the mass appeal and presence of *Spacewar!* within laboratories around the world would eventually influence the creation of the first commercially available arcade video game title ever released, *Computer Space* (1971); a game which set precedent for not only arcade video games as a whole, but also the ways in which they utilized audio as a component of their design.

Following the completion of his electrical engineering degree at the University of Utah in 1968, Nolan Bushnell, *Computer Space* co-creator and eventual founder of Atari Inc., moved to California to accept employment at Ampex, one of the leading audio/video hardware manufacturers of the time (Edwards). It was at Ampex that Bushnell was introduced to the other half of the *Computer Space* development team, Ted Dabney, who together with Bushnell would be inspired to develop *Computer Space* from their extensive time spent playing *Spacewar!* at Stanford University's Artificial Intelligence Laboratory (Nolan). The duo's initial goal was to develop a coin-operated version of *Spacewar!*, which ultimately would not come to fruition, due to the exorbitant cost of appropriately powerful microcomputers at the time (Edwards). Instead, the team moved forward with the project through the development of transistor-transistor logic (TTL) circuits, which allowed for video generation and manipulation utilizing discrete components including integrated circuits, transistors, and diodes (Lowood). It was through the use of these component based TTL circuits that Dabney was able to develop a system which provided both video and sound effects for the gameplay without the necessity to deploy a microcomputer.

The exact means by which Dabney was able to generate the sound for *Computer Space* resulted of his use of a 6V Zener diode to regulate voltage in the video circuits. Dabney was aware that Zener diodes produce pink noise during operation, and through the addition an "amplifier and an integrator that charged up and decayed to fade the volume out," he was able to utilize the pink noise as the basis for the sound effects which accompanied the presence of *Computer Space* in arcades around the country ("The Making of"). Dabney's method of sound generation for *Computer Space* was a unique result of the team's reliance on the use of discrete component-based TTL circuits to generate video signals which could be variably controlled by the input provided by the player. Without the use of a microcomputer, with which the game could be coded in an assembly programing language, and allow for specific sound generation on a software level, Nolan Bushnell and Ted Dabney's first entry into the coin arcade video game was dependent on tapping into the existing TTL circuit to find sounds which would prove worthwhile for the game; an approach which would also be used in the development of Bushnell's next coin operated arcade game, Atari's first release as a company, and the game that truly brought video game into the public mainstream: *Pong*.

The development of *Pong* is of instrumental importance in the history of video games, as well as in the focused study of their audio components, for the very reason that it was the first arcade video game with sound effects to break through the commercial barrier that video games had been held behind thus far; resulting in its gameplay, visuals, and sounds becoming synonymous with the advent of video games in the hearts and minds of millions of individuals around the world. In the year following the release of *Computer Space*, Nolan Bushnell was introduced to the Magnavox Odyssey; the silent system which I previously mentioned above as being the first home gaming console to ever be released. It was at a Magnavox sponsored demonstration of the Odyssey in May 1972 that Bushnell played the system's *Tennis* game, which would inspire his aspiration to develop something of a similar vein (Winter, ”Atari PONG"). Shortly after releasing *Computer Space*, Bushnell and co-founder Ted Dabney hired a fellow coworker from Ampex named Allan Alcorn to help with the development of their next arcade title, which would, of course, be *Pong*. According to David Winter, writer and operator of the web-based early video game archive, “Pong-Story,” the title of *Pong* was chosen as its meaning, “a hollow, ringing sound,” was exactly what Bushnell wanted to hear during gameplay (“Atari PONG”).

I am hesitant to believe that the title of *Pong* is anything more than a shortened form of “Ping Pong,” as the game is obviously inspired by Table Tennis, a sport of which “Ping Pong” is a common moniker for, though I believe it is apparent to anyone who has heard the classic sound effects of Atari’s first success that they can be accurately described as “pong-like.” However, Bushnell’s true desire for the sound of *Pong* has been described by Allan Alcorn himself in his *Engineering and Technology History Wiki* contribution, “First-Hand: The Development of Pong: Early Days of Atari and the Video Game Industry." As Alcorn states in this first hand account:

> Nolan said that it had to have sound. But I was already over budget. I was finished with it except for the sound. So Nolan said, “I want to have the sound cheering people when you make a score.” And Ted said, “I want it to have boos and hisses.” I didn't know how to [sic] either one with digital components. I was already over budget. Then one afternoon I poked around the synch generator and found tones that were already there, and that's where the sounds came from, and that became the sounds. Articles have been written about how wonderful the sounds were. And I told Nolan, “if you don't like the sounds, you go do something

> better, but I'm only doing that." And so that's how the sounds got out. But it was because of cost… I was over budget.

Given the fact that *Pong* was developed with the same TTL based circuitry as *Computer Space*, there was no simple way for Alcorn to create audio for the game that would fulfill the expectations of Bushnell. Had the game been developed with a CPU as a primary component, Alcorn would have been granted more freedom in the generation of sound content for the game, most likely in the form of simple digital sound synthesis. And, in addition to being able to synthesize sound effects more in-line with the desires of Bushnell, Alcorn would have also been able to design the synthesis of the audio in a way which would allow it to be completely reactive to the gameplay from a programming perspective; though this lack of programmability proved to be of little issue to Acorn.

While the lack of a capable programable sound generation device severely limited Alcorn's ability to create the audio that Bushnell and Dabney both envisioned for the game, much in the same way that Dabney was able to route current away from the Zener diode of his TTL circuit though an amplifier to produce the sound for *Computer Space*, Alcorn's use of the video synch generator as the source of sound generation resulted in sound effects that were completely reliant upon changes in the current that were reflective of gameplay changes. As the amplifier was connected directly to a point on the video synch generator, all changes in sound content were a direct result of changes to the video output. What I found most interesting in this regard is the fact that the sounds that are heard when playing the classic arcade game were "chosen" by Alcorn for the simple reason that they were the most appropriate tones available to be amplified from the synch generator for any given visual change in the direction that the ball is traveling, including: bouncing off of either players paddle, the boundaries of the play area, or leaving the play area entirely. The difference in pitch that is heard between the sound effect that is rendered when the ball hits the paddle, versus when the ball hits the boundaries of the play area is a direct result of the differences in the electronic signal being generated by the synch generators. The pitches that accompany each action were not freely chosen by Alcorn, yet, they work exceptionally well in the context of the game; so well, in fact, that sound effects of *Pong* have become iconic in their own right, with many recognizing the simple “beeps and bloops” as a true progenitor of video game audio into the modern era.

<u>The Dawn of the Home Console</u>

It should be obvious that *Computer Space* and *Pong* were not the only arcade video games which utilized audio. Throughout the years following the release of *Pong*, arcade video games experienced a major increase in popularity, with many titles relying on flashy visuals and catchy sound effects to draw the attention of potential players within arcades around the world. During this period, many arcade video game manufacturers utilized TTL circuit technology for their games, as it remained significantly less expensive than the use of programable microcomputers of the time. The use of TTL circuits in Atari hardware changed in 1975 with the initial sale of radically improved *Pong* systems, which relied on integrated circuits (ICs) as a simplification of the TTL technology to provide the paddle and ball style gameplay that the game had made popular just two years prior (Winter, “PONG in a Chip”). Ultimately, there was little need to bolster Atari’s arcade market with these new IC based machines, as the pure TTL based units were still maintaining a high level of popularity within arcade establishments. What Atari had hoped to do with their new IC based systems was to create a competitive home video game economy; one which would replace the previous market, which had been widely ruled by the Magnavox Odyssey since its release in 1972.

Recognizing the demand for "Pong in a Chip" systems, many IC fabricators began releasing their own chips which ran clones of *Pong*, as well as other similarly designed games, in hopes of cashing in on the newest fad in home entertainment. This resulted in a flood of home "Pong in a Chip" systems being released by numerous manufacturers, who included chips which were being produced by a small number of fabricators at the time. This meant that while many of these systems were of different manufacturers and were marketed as being unique, they were essentially identical from a gameplay perspective. And as these systems utilized what essentially amounted to a condensed version of TTL technology, they were still unable to programed for audio purposes.

The prevalence of "Pong in a Chip" systems within the video game market was challenged with the release of the Atari 2600 (also known as the Video Computer System, or "VCS") in 1977. While not the first home console to utilize a microprocessor and ROM cartridges for game program storage, as the Fairchild Channel F had been previously released in 1976, the Atari 2600 held 80% of the home video game market share during the period, and can be credited with introducing the largest user base to the audio capabilities that accompanied the use of a programable processor unit in a home gaming console (Torres). The most notable benefit of the use of programed audio in a home console came in the form of continuous musical accompaniment during gameplay. As the introductory source of programed video game music to millions of people in the 1970s, the Atari 2600 cemented the initial standard by which all future home video game audio would be compared. Though, that is not to say that the 2600 set a particularly high standard. Marred by the technical process by which its audio was generated, resulting in pitches that could not be produced in tune with each other, the 2600 suffered from issues of dissonance when playing back harmonically supported melodies.

This issue of improper tuning was due to the method used by the 2600's sound card, the Television Interface Adapter (TIA), to synthesize the music on playback. According to Karen Collins, author of *Game Sound: An Introduction to the History, Theory, and Practice of Video Game Music and Sound Design,* "the awkward tuning on the VCS was due to the TIA's 5-bit pseudo-random frequency divider, capable of dividing a base frequency of 30kHz by 32 values (21). Without going into terrible detail as to why this results in out of tune harmonic content, I can simply say that the equal temperament tuning system that most people, particularly those with "western ears," are used to hearing in modern music cannot exist in a system such as that utilized by the TIA.

The TIA also suffered from the limitation of only being capable of generating two channels of audio at any given time. This meant that the music heard during game play could only be as complex as a melody and single harmonic line, and in the event that a reactive sound effect was triggered by a gameplay action, the music would often lose its harmony, or drop out entirely, to allow the effect to be produced.

While the Atari 2600 was the most popular console of the late 1970s and early 1980s, it received some serious competition, especially in regards to the quality of its audio. The most notable of these competitor systems would be Mattel's Intellivision, which was released three years after the Atari 2600, in 1979. The primary ways in which the Intellivision far exceeded the audio capabilities of the Atari 2600 lay in its use of a General Instruments AY-3-8910 three channel sound card, which was able to create three voice harmonic musical content, as well as generate game play sound effects while maintaining two voice harmonic music (Barton, Loguidice). This is an obvious step up from the limitations held in place by the two channel TIA found in the Atari 2600. Additionally, the AY-3-8910 was capable of generating tonally correct harmony, which left players with an immediate recognition of improved sound quality from the console.

Arguably, the most notable sound capability of the Intellivision came in the form the Intellivoice add-on, a voice synthesis device which was released in 1982. Through the addition of the Intellivoice, game programers were able to implement synthesized speech as an additional sound source within their games. Throughout the Intellivoice's lifecycle, only five games were developed to utilize the peripheral, though, these games existed as some of the console's most distinct and sophisticated titles (Barton, Loguidice). The power of the Intellivoice add on can be seen most prevalently in the 1983 Intellivision title, *World Series Major League Baseball*, which is impressive for being "one of the first multi-angle baseball games, with its fast-paced gameplay based on real statistics, play-by-play with the Intellivoice, and the ability to load and save games and lineups from cassette" (Barton, Loguidice). The ability for the developers to program reactive play-by-play speech to accompany gameplay was ahead of its time for a home video game console in 1983, and sets the Intellivision apart as one of the most important consoles of its time in the survey of video game audio.

Heading into the mid-1980s, the North American home video market began to suffer devastating losses, due to over-saturation of consoles and games being released by large numbers of manufacturers and developers, as well as the constant release of less-than-adequate ports of beloved arcade game. There were simply too many choices available, and too much disappointment with the quality of titles, with supply severely outclassing demand (Lambie). Though the market was in the midst of decline at the time of its release in 1985, the Nintendo Entertainment System (NES) proved to be one of the most notable consoles of all time, especially in regards to its audio performance. Building from the precedent set by the ROM based programable consoles of the late 1970s, the NES was released on a marketing platform reliant on game sales. In my opinion, the immense success of the NES in the mid-to-late 1980s was responsible for the revitalization of the home console industry into, and throughout, the 1990s, and this success can be owed to the quality of the first party titles being released by Nintendo and its subsidiary developers. Gone were the days of sub-par arcade ports; the NES's catalog was chock full of interesting and attention demanding titles which could only be experienced at home, on the hardware. In many ways, the success of these titles are a result of their entertaining gameplay, interesting characters, and their use of some of the best music and sound effects of any early gaming console.

The quality of the audio produced by the NES is a result of its programable sound generator (PSG), the RICOH RP2A03 Sound Processor, which boasted an impressive five channels of concurrent audio generation, including two pulse waves, one triangle wave, a noise generator, and one channel capable of playing back 7-bit pulse code modulation (PCM) digital audio files (“RP2A03”). Compared to the other popular consoles of the NES’s era, including the Atari 2600 and Intellivision, the NES is an audio workhorse. The use of five channels of audio allowed developers to program rich harmonic music with room for effects, and PCM playback allowed for more realistic speech production than that which was available with the Intellivoice, though there was a huge limitation on the length and frequency of PCM playback inherent to the small storage capacity available on the NES cartridges. Nonetheless, many of the most recognizable video game soundtracks and musical themes of all time, including those which accompany Nintendo’s iconic *Super Mario* and *Legend of Zelda* series, got their start as programed music on the NES.

Expanding Audio Capabilities in the Fourth Generation

Through the critical and commercial success of the NES, the home console market began to rebuild, with many manufacturers and developers showing less hesitation to enter the industry that was once on the verge of collapse. One such company who brazenly entered the recovering video game market was SEGA, who had introduced their competitor to the NES, the Master System, to the North American market in 1986. Unfortunately, the Master System was rewarded little success in the shadow of the market dominating NES. However, this did not deter SEGA. Following years of research and development, SEGA went on to initiate the fourth generation of home video game consoles with the release of the Genesis in 1989; one of the most important consoles in regards to games audio, due to its use of frequency modulation (FM) synthesis as the basis for its sound capabilities.

Though still a relatively new technology at the time of the Genesis's release in 1989, FM synthesis had already been lauded for its ability to produce near realistic, and complex, sounds from combinations of simple wave forms. Used extensively in Yamaha's early synthesizers, FM synthesis yielded an interesting sound that would find its way into many synthesizer and electronics-rooted records of the 1970s and 1980s. SEGA's inclusion of a Yamaha designed YM2612 six channel FM synthesis chip, as well as a four channel Texas Instruments SN76489A PSG, allowed the Genesis ten channels of rich audio support for gameplay ("Sega Genesis Technical Manual"; "SN76489"). In many cases, the use of the FM synth created soundtracks and sound effects that sounded nearer to real world instruments and sounds than any other console was capable of; putting the Genesis a step ahead of its competition in the context of audio. Additionally, the Texas Instruments PSG also allowed for the generation of pulse waves, white noise, and low bit-depth PCM files that were already commonly associated with video games of the time. Ultimately, the Genesis was one of the few consoles to rely on FM synthesis as a means for audio generation, and as a result, still remains recognizable as one of the most unique sounding pieces of video game hardware ever developed.

Following their commercial success with the NES, Nintendo released the Super Nintendo Entertainment System (SNES) in 1990. The SNES holds the position of being the most popular of the 16-bit fourth generation consoles, outselling the SEGA Genesis, and remaining popular well into the next generation of consoles that would follow. In similarity to the NES, the SNES owes much of its popularity to the first party content which was continuously being developed by Nintendo, as it was built upon the success of franchises that saw their roots on the NES. Through the truly advanced video and sound capabilities of the SNES, many of these franchises saw a rebirth that can be well recognized in their current iterations today. While the NES may have introduced the characters, gameplay experiences, and music that is so readily associated with Nintendo's most popular intellectual properties, their home truly is with the SNES.

In my opinion, the most important effect that the SNES had on these franchises was in helping to flesh out their musical identities; something that was made possible though Nintendo's inclusion of the Sony designed SPC700 audio processor, also known as the S-SMP. It was though the use of the S-SMP that the SNES was able to generate some of the most realistic sounds of the era, in many cases outperforming the Genesis due to the S-SMP's ability to playback PCM files at near CD-quality. However, this was not without some pretty severe limitations. As explained in the *Games Are Evil* article, "The Rich Tones of FM Synthesis":

> While FM was cutting edge for the Genesis in 1989, it was quickly out-shined a couple of years later by the Super Nintendo. Nintendo's 16-bit console was late to the party for many reasons, but one of those reasons was to make sure that the SNES used the best technology available on the market. Nintendo's SPC700 audio processor (from Sony) was marvelous for the time. It provided real sample playback. Composers would manipulate those samples in many different ways. One problem however was the lack of available sound memory. As a result, the sounds were often either incredibly short, which led to a warbly sound when

> sustained during playback (the strings in *Star Fox, Yoshi's Island*, and *Final Fantasy VI* suffer from this) or sampled at a high pitch then slowed down in software to save space and include more samples. This led to the trademark SNES "muffled sound."

Though the SNES was restricted in its use of raw digital audio files as a result of memory deficiencies, using the S-SMP's ability to produce and mix eight simultaneous voices, pan voices, generate ADSR volume envelopes, apply echo with an eight-tap finite impulse response (FIR), and use sound as a noise source, Nintendo's in-house SNES developers, and their contemporaries, were able to create experiences that remain iconic to this day ("Nintendo S-SMP").

32-bit Systems and the Beginnings of CD Quality Games Audio

As a result of the development of the first 32-bit games consoles, the mid-1990s brought many major advances to the quality of home video games across all fronts. The visuals reached a level of realism unseen prior, with the audio improving equally in tandem. These improvements were incredibly dramatic, leaving many players to question where the video game experience could possibly go from there. The forerunner in the development of these, for the time, remarkably immersive experiences was the original Sony Playstation, released in 1994. As the first commercially successful console to feature the use of an optical disc as the program storage medium, the Playstation was capable of utilizing this technology to enable unparalleled sound in the form of full 16-bit 44.1kHz CD-quality audio playback. For the first time, not only could game developers and composers create music and sound effects that accompanied the game play with incredible depth, but also utilize playback of digital recordings in their games. Games such as *Wipeout* (1995) and *Tony Hawk's Pro Skater* (1999) are still known today as examples of the ways in which the Playstation pioneered the use of popular styles of recorded music as effective musical accompaniment in modern games.

Aside from its CD-quality digital audio playback capabilities, the Playstation is also important in games audio for its sound processing technology. Being based on 32-bit architecture, the Playstation, naturally, has much more processing power afforded to its production of audio. The console was capable of producing 24 channels of 16-bit PCM files, as well as midi instrument playback ("Playstation Technical Specifications”). This means that the console was capable of generating synthesized and sample based music as well as PCM based sounds. In addition to its impressive channel count, the Playstation was able to process audio through pitch modulation, ADSR volume envelope, looping, and digital reverb ("Playstation Technical Specifications”). These tools granted audio programers a previously unheard of level of mix control within their games. The technical audio capabilities of the Playstation can be easily recognized in the 1996 title, *PaRappa the Rapper*, which was the first game to ever win the Academy of Interactive Arts and Sciences's award for “Outstanding Achievement in Sound and Music” in 1998, two full years after its release (“1998 Awards Category Details”).

Reactive and Interactive Mixing in Modern Games Development

Now that we have covered the breadth of video game audio capabilities, surveying the progression from the earliest games of the mid-20th century all the way up until the beginnings of modern CD-quality games audio, we can begin to look at how developers are currently tackling the challenge of creating the dynamic, interactive, audio experiences that gamers are enjoying today. As technical capabilities improved following the Playstation and its contemporaries, with the release of sixth generation consoles, such as the Sony Playstation 2, Nintendo Gamecube, and Microsoft Xbox, the role of mixing the audio for video games changed dramatically. Where it was once the audio programer's job to provide a system that played back either programed or prerecorded music and triggered sound effects reactive to player actions within a static mix, the increase in concurrent channel counts and processing capacity with these newer consoles left developers looking for a way to handle the audio mix in a dynamic way, reactive to not only triggered actions made my the player, but also to the player's presence in the world of the game. This lead to the development of dynamic audio mixing engines, which handle many mixing and audio asset instantiating decisions dynamically, and in real time, entirely in relation to the input of the player. Using such systems, the creation and integration of music and sounds has become both more involved, and more interactive than ever before. These improvements in capabilities have allowed the act of interactive mixing within games to become more pronounced, and with so many concurrent sources of sound (background music, dialogue, ambient sounds, etc.), more important than ever.

It is important to discuss the importance of the processes that go into mixing audio for video games. According to Garry Taylor, Audio Director at Sony Computer Entertainment Worldwide, "...the audio assets that go into a game are only 50% of the complete experience. The other 50% is down to implementation of those assets, with a good mix being a large part of that." Throughout his 2010 Game Developer Conference keynote from which this quote was taken, Taylor discusses the types of processes that go into creating a mix that fulfills the other 50% of the complete interactive audio experience. Mixing for video games is much the same as mixing for film and television, with the end goal being to create a listening experience that is clear and defined. In many ways, this goal is achieved as a result of the same process by which the mixing engineer for these other non-interactive mediums performs their job. To create a clear and balanced mix, the engineer must make sure that every sound source has a defined location within both the frequency spectrum and the stereo, or surround, image. Taylor sums the processes of mixing for video games into two distinct categories: active and passive mixing. Both methods are still used today, though, through the use of dynamic mixing engines, active mixing is beginning to take hold as the standard for the incredibly immersive experiences modern gamers desire.

Active mixing is accomplished by creating a system that allows actions and events within the game to alter the mix according to their function and importance. Examples of this process as described by Taylor are, “the recalling of a set of volumes for a group of sounds, a snapshot, triggered by an in-game event, or for example the tinnitus effect when grenade goes off close to you, in a first person shooter, where all the sound is filtered, except the ringing in your ears effect." This description accurately outlines some benefits of the active mixing process. Due to the fact that the in-game audio can adjust fluidly to match player actions, active mixing is quickly becoming the standard process for mixing audio in the video game industry.

On the other end of the Taylor’s spectrum we have passive mixing. Taylor describes passive mixing as similar to the standard way in which an engineer would mix music or film audio. Passive mixing is accomplished by setting dynamics processors to static configurations, allowing other sources of audio to trigger a change in the mix through various routing techniques. This can be as simple as compressor pulling back the music when dialogue enters a side-chain input. While this may sound similar to the goal of active mixing, in that the mix is being altered by changes in the other sounds present in the mix, there is a distinct difference. With passive mixing, the changes to the mix are only based on what is coming out of specific channels which are predetermined to be routed though side-chain compressors or other dynamics processors. The mix never adjusts to what is happening on screen, only to other audio sources. In the case of differentiating active and passive mixing, we must recognize that this passive process of side-chained ducking is naturally included within the active mixing ideal. Taylor describes his perfect setup as, “a combination of both active and passive systems," for this very reason. Though active mixing naturally includes passive processes, the same can not be said in reverse. Passive mixing does not feature any techniques that fall in line with the ideal of active mixing, and is typically handled using the static mixing techniques that have been utilized for game development since the advent of programable audio.

As far as the techniques involved in active mixing go, there are a variety of different approaches being taken. Some developers include proprietary systems into their games that allow snapshots of the mixer to be loaded on the fly when certain events are triggered, while others choose to integrate commercial systems such as Firelight Technologies’s FMOD, AudioKinetic’s Wwise, and D.I.C.E.’s High Dynamic

Range (HDR) which dynamically mix the audio based on various parameters set by the mixing engineer during the production of the game.

The technique of utilizing in-game actions and events to trigger mixer snapshots is very easy to understand. The mixing engineer sets different mixes for different locations and events in the game and saves these as snapshots. These are then loaded back into the games internal mixer when individually triggered by the appropriate in game sequence, or player input. In most instances, the transition between mixes is not immediate, which enables the change in sound to go almost unnoticed. This allows for different levels of foley and ambient sounds to build as many different soundscapes as the game might require. Rob Bridgett, audio director at Radical Entertainment, explains in his article, “The Future of Game Audio - Is Interactive Mixing The Key," that mixing engineers must take their goals from film mixing, but also need to expand upon this using automated systems in order to overcome the non-linear nature of most games. Bridgett implemented a proprietary snapshot technology for Radical’s 2006 game *Scarface: The World is Yours*, which allowed Bridgett and his team to mix various instances of unique audio within the game individually, based upon what sounds were present at that time. At Radical games, Bridgett found it helpful to divide snapshots into three groups based on the function that the snapshot serves. These snapshot categories were: Default Snapshots, Multiplier or 'ducking' Snapshots (several of these can be installed at any one time, and when added to a default snapshot, these create a ducking effect), and Override Snapshots or event snapshots (only one can be loaded at a time, as these completely override any default or ducking snapshots that may be installed for the duration of an event) ("The Future of Game Audio”). With the snapshot method, the audio in games can consist of multiples of individual mixes. According to Bridgett, *Scarface: The World is Yours* contained roughly 150 individual mixer snapshots, each used in specifically outlined instances within the game ("The Game Audio Mixing Revolution"). In 2006, when *Scarface: The World is Yours* was developed, this snapshot technology was more advanced than the mixing systems most developers were using at the time. Over the next few years snapshot mixing became common within the industry, however, it is now being replaced by even more fluid and dynamic mixing technologies.

While mixer snapshots allow for active changes in the mix based on actions in the game, they have a major shortcoming when it comes to handling the demanding audio that most modern games are requiring. With games sometimes featuring huge amounts of concurrent sounds, mixing screenshots can be limited in many ways; channel count and dynamic headroom just to name a few. This is where the automated mixing systems (FMOD, Wwise, HDR) come into play. These systems adjust the importance of audio files in the context of the current situation within the game. Yes, snapshot mixing can adjust the importance of certain sounds, but it does not react to level differences on the fly, as all changes are predetermined by the mixing engineer. D.I.C.E.'s HDR system, on the other hand, was designed specifically to adjust levels in real time, based on the importance of the sounds. According to Garry Taylor, the system works by creating a window that has a defined dB range to cover the dynamic range of a given system governed by the playback method chosen by the user; headphones, internal television speakers, home cinema system, etc. Templates are then created which contain the samples as well as control data, including loudness data for each sound. The system reacts to the introduction of really loud sounds by shifting the window up to allow the newly introduced loud sounds to be contained at the top of the window. Sounds that fall below the new bottom of the window are then attenuated. When the loud sound fades away, the window shifts back down bringing the quieter sounds back into the mix. This system can handle individual track level changes without the need for mixer snapshots, as most changes in the overall mix come from the addition and removal of sounds. Due to the shifting nature of the window, the dynamic range remains static throughout the adjustments being made to the mix.

Along with HDR, there are various other active mixing systems available to the mixing engineer. Taylor describes a category of these as “self-aware systems." The use of self-aware systems allows for the mix to expand upon what was predetermined by the engineer. This category of systems, of which FMOD and Wwise fall under, are integrated into the playback engine within the game and are constantly aware of what function each sound holds in the mix. The adjustments to the mix are not triggered by events like snapshots, or by the introduction of sounds, like HDR. The self-aware systems take information provided to them and function only to use this information to make necessary adjustments. According to Taylor, “it’s now possible for us to go one stage further and have systems that are aware of what they’re actually outputting, and to make passive mix decisions accordingly.” He continues by explaining that, “you can store spectral information as metadata about each audio file in your game, and you know how loud each sound is being played and where it is in the 3D world, the system can know exactly what is coming out of each speaker at any moment.”

The benefits to this level of autonomous control might not be obvious, but they are substantial. The HDR system can dynamically handle ducking and level changes based on incoming audio sources. This is an impressive automated function. However, it is limited to only adjusting levels. On the other hand, self-aware systems not only make automatic changes to levels, but, due to their knowledge of sound location in 3D space and the spectral components of each sample, are able to make adjustments to EQ and dynamics processing in order to actually allow more clarity for new sounds, without only adjusting levels and removing quieter sources. This is a huge advancement from the simple triggering of pre-made mixes used in the snapshot based systems. Additionally, through the use of various processing techniques, such as applying convolution based reverb to all sounds in a similar space and changing parameters of the effect based on the location of these sound within the space, dynamic mixing engines can create startlingly convincing aural environments for the player's ears to explore. When developers add in support for the use of five, seven, and even nine channel surround sound systems, or include head-related transfer functions (HRTFs), which are mathematical formulas that are used to process the audio being fed to a set of standard headphones to make the stereo image sound as if it is entirely surrounding the listener, the dynamic mixing engines are able to accurately position the sounds within the 360 degree sound image in a way that envelopes the player in sound, as if they were physically located within the world of the game. As the player moves through the in-game space, the processing and positioning of sounds adjusts to create the most realistic representation of the soundscape as possible; which in itself is one of the most effective ways of promoting total player immersion.

The use of active mixing within game sound design essentially creates a totally dynamic mixing environment within the world of the game, in which all changes to the sonic characteristics of sounds can be modified in realtime, reacting to the players actions and in-game events. This process is especially seen with the use of self-aware systems. As I mentioned before, most developers use a combination of active and passive mixing due to the complex nature of mixing audio in general. In line with this, most developers also use a combination of different methods of active mixing in order to create the soundscapes we hear today. Some may say that the development of completely automated systems can remove the job of the mixing engineer in games development entirely, but this will never be true. The use of these systems are the same as any other tool in the mixing engineer's arsenal. These systems allow the developers to overcome problems inherent with attempting to mix a non-linear format in the way one might mix linear sounds, like music or film. The mixing engineer will always be necessary, using these systems to create the best possible sound for the consumer to enjoy.

Conclusion

I believe it is apparent that video games have experienced a huge transformation in sound capabilities from their very humble beginnings in research laboratories as technical demonstrations of the most powerful technology of the time, to the modern systems currently found in homes around the world, which are capable of creating intense, dynamic, and realistic audio experiences that far exceed anything that was ever possible in the earliest days of video games. The progression from the simplistic circuit tapped audio of *Computer Space* and *Pong* in the 1970s, to the self-aware dynamic mixing engines of today is indicative of the nature in which all technology has developed and evolved over the same period, and while some may feel that the rapid pace of technological revolution that has taken place over the last century is coming to a halt, I believe that in looking at the last 60 years of video game audio development, regardless of it specificity, one can easily see the sheer power of human ingenuity and our drive to build and develop new technologies. It is with our propensity to design technological evolution that video game audio has gotten to the impressive point of technical capability that is exists in today. And it will be with the same propensity that in short time we will experience the next major evolution in games sound. After all, video games are still rapidly growing as as one of the most influential and widely consumed media forms of all time. There should be no doubt that the development of new gaming technologies, for audio or otherwise, will continue well into the future.

Works Cited

"1998 Awards Category Details - Outstanding Achievement in Sound and Music." *Interactive*. Academy of Interactive Arts and Sciences. Web. 13 May 2015.

“2014 Essential Facts About the Computer and Video Game Industry." Entertainment Software Association. Oct. 2014. Web. 1 Apr. 2015.

Alcorn, Allan. “First-Hand: The Development of Pong: Early Days of Atari and the Video Game Industry." *Engineering and Technology History Wiki*. 12 Jan. 2015. Web. 10 May 2015.

Bateman, Chris. "Meet Bertie the Brain, the World's First Arcade Game, Built in Toronto." *Spacing Toronto*. 13 Aug. 2014. Web. 08 Apr. 2015.

Barton, Matt, and Bill Loguidice. "A History of Gaming Platforms: Mattel Intellivision." *Gamasutra*. Web. 12 May 2015.

Bridgett, Rob. "The Future Of Game Audio - Is Interactive Mixing The Key?" *Gamasutra.* 13 May 2009. Web. 9 Oct. 2013.

Bridgett, Rob. "The Game Audio Mixing Revolution." *Gamasutra.* 18 June 2009. Web. 9 Oct. 2013.

Collins, Karen. *Game Sound: An Introduction to the History, Theory, and Practice of Video Game Music and Sound Design*. Cambridge, MA: MIT, 2008. Print.

Edwards, Benj. "Computer Space and the Dawn of the Arcade Video Game." *Technologizer*. 12 Dec. 2011. Web. 20 Apr. 2015.

"First Video Game?" *YouTube*. Brookhaven National Laboratory, 21 Oct. 2008. Web. 17 Apr. 2015.

Hatzithomas, Ion. "Good Deal Games - Classic Videogame Games ARTICLE - Home Video Game Console History." *Home Video Game Console History*. Good Deal Games, Web. 08 Apr. 2015.

Lambie, Ryan. "The 1983 Videogame Crash: Could It Happen Again?" *Den of Geek*. 20 Feb. 2013. Web. 12 May 2015.

Lowood, Henry. "Videogames in Computer Space: The Complex History of Pong." *IEEE Annals of the History of Computing*. IEEE Computer Society. Web. 29 Apr. 2015.

"NIMROD: The First Special Purpose Digital Computer Designed to Play a Game (May 5, 1951)." *History of Information*. Web. 08 Apr. 2015.

"Nintendo S-SMP." *Wikipedia.* Wikimedia Foundation. Web. 12 May 2015.

"Nolan Bushnell and Digging up Spacewar! - Atari Inc. - Business Is Fun." Web log post. *Atari Inc. Business Is Fun*. 11 Jan. 2014. Web. 02 May 2015.

Peddie, Jon. *The History of Visual Magic in Computers: How Beautiful Images Are Made in CAD, 3D, VR and AR*. New York City: Springer, 2013. Print.

"Playstation Technical Specifications." *Wikipedia*. Wikimedia Foundation. Web. 13 May 2015.

"RP2A03." *Videogame Music Preservation Foundation Wiki*. Web. 12 May 2015.

"Sega Genesis Technical Manual - YM2612 Section." *Maxim's World of Stuff*. Web. 12 May 2015.

"SN76489." *Sega Retro*. Web. 12 May 2015.

Taylor, Garry. "All In The Mix - The Importance Of Real-Time Mixing In Video Games." Web log post. *Blessed Are The Noisemakers: A Blog on the Creative and Technical Side of Video Game Audio.* 20 Feb. 2012. Web. 9 Oct. 2013.

"The First Video Game?" *BNL*. Brookhaven National Laboratory. Web. 17 Apr. 2015.

"The Making Of Computer Space." *NowGamer*. 07 Feb. 2012. Web. 03 May 2015.

"The Rich Tones of FM Synthesis." *Games Are Evil*. Web. 12 May 2015.

Torres, Gabriel. "Inside the Atari 2600." *Hardware Secrets*. 12 May 2012. Web. 13 May 2015.

"Video Game." Def. 1. Video Game: Definition of Video Game in Oxford Dictionary. *Oxford Dictionaries*. Web. 8 Apr. 2015.

Winter, David. "Atari PONG - First Steps." *Pong-Story*. Web. 3 May 2015.

Winter, David. "PONG in a Chip." *Pong-Story*. Web. 12 May 2015.