CHAPTER 4

Press Reset: Video Game Music Comes of Age

In 1988 Williams's *NARC*, the first 32-bit arcade game was released, but the arcades—which had seen a decline during the 16-bit era—had now almost disappeared. Even though earlier arcade games had typically been far more advanced in sound technology than home games, with the arrival of 32-bit machines and CD-ROM technology, home consoles and computers soon overtook the arcades. At a time of decreasing revenues, many coin-op manufacturers kept to older, more affordable sound chips. Although graphics continued to improve, numerous arcade games released by the smaller companies throughout the 1990s continued to use the same sound technology that had been around for a decade or more.[1]

By the early 1990s, most home computers had FM soundcards supporting MIDI, but many of these soundcards were cheaply produced, and the FM synthesis made MIDI music sound disappointing in the face of CD releases. When CD-ROMs became popular, MIDI in games was for a large part abandoned in favor of more realistic sounds. Popular early CD-ROM titles like *7th Guest* (Trilobite, 1992, music by George Sanger) were released with impressive high-resolution graphics and live recorded music.[2] Since the audio was not reliant on a soundcard's synthesis, CD-ROM technology ensured that composers and sound designers could not only know how the audio would sound on most consumer configurations; they could also now record sound effects, live instruments, vocals, and in-game dialogue. The downside of the CD-ROM technology was that the discs could hold only a maximum of 72 minutes of uncompressed *Redbook* (CD) audio. With a game included on the disc, audio still had to fight for space. As various compression technologies were developed (such as MP3), however, much less data was required to store audio, and game companies began incorporating more compressed audio into their games.

Another important advance in game audio in the 1990s was the ongoing evolution of three-dimensional (3D) or *surround* sound. Surround sound gives the listener the perception that the sounds are emanating from a three-dimensional space. It is likely that, other than in movie cinemas, it was in the arcades that many gamers first experienced surround sound, although it was nearly exclusively used in racing and simulator games. The first PC soundcard to support surround sound was Diamond Monster Sound, in 1997. Parameters such as a room's size and acoustic properties could be programmed into a game, which would initiate filters and effects to simulate the space (see Miller 1999). Surround sound would become an important element in game design in the following decade, as Matthew Lee Johnston of Microsoft explained: ''Traditional stereo has been used to localize a sound in the player's forward visual field. What 3D audio adds is the ability to localize the sound behind the player, which is arguably way more important, since the sound is usually the only way to provide the player with feedback about what's going on behind them. Some games use maps and have 'rear view' options, or even let you pan your visual field around to look, but using 3D audio to position an object behind the player is not only more immediate and instinctual, but it allows the player to focus simultaneously on the fore and aft perspective (cited in Miller 1999).

In order to create a three-dimensional atmosphere, the system must position the sound source and the listener in a three-dimensional environment, define the environment in terms of dimensions and acoustic properties, and synchronize the sound to graphics. A dynamic 3D audio system must also process this in real time, based on the location of the player's character. There were many different technologies developed in the 1990s that called themselves ''3D'' but which varied greatly in quality, such as *spatialized 3D* or *virtual surround* (Warwick 1998). Of particular relevance to games was *positional audio*, which used signal processing to locate a single sound in a specific location in a 3D space. With 3D positional audio, sound objects in a virtual space could maintain their location or path of motion while the gamer's character moved about (ibid.). Volume was then adjusted depending on the distance to the player's character.

To effectively implement surround sound technology into games, new software developments were required. Beginning with *Windows 95*, the Microsoft *Windows* platform was packaged with DirectX, a series of multimedia application programming interfaces (APIs) that improved the speed with which sound and graphics cards could communicate.[3] Quickly becoming the standard, DirectX allowed games programmers to access ''specialized hardware features without having to write hardware-specific code'' (Warwick 1998). In other words, the DirectX interface bridged software and hardware, allowing for better control of sound mixing and output. One element of DirectX, called DirectMusic, was a particular advance for music in games. DirectMusic overhauled the older MIDI protocols by offering the industry-ratified Downloadable Sounds Level 1 (DLS-1) specifications, providing over one thousand channels, with better timing mechanisms, and real-time control (see Hays 1998).[4] DirectMusic opened up MIDI to the possibility of higher-fidelity wavetable synthesis and sampling, while at the same time it allowed MIDI to have more extensible controls. Now, no matter what soundcard a player had on his or her personal computer, the results could be much more predictable.

Gaming on personal computers had by the 1990s developed into a rapidly evolving industry, with far too many variants to discuss here. I will, therefore, focus on three examples of the most popular games that exemplify different approaches and technologies: a puzzle adventure game (*Myst*), a first-person shooter (*Doom*), and a simulation game (*The Sims*). As others have noted, ''in 1993 [the popularity of] two games signalled that something new was happening in computer gaming. They were *Doom* and *Myst*'' (Kline, Dyer-Witherford, and de Peuter 2003, p. 113). *The Sims* came later, but it quickly became one of the best-selling computer games of all time.

Surround sound conveniently arose at about the same time as a new genre in games, the first-person 3D shooter (often now referred to as *FPS*).[5] 3D shooters were born with the release of *Wolfenstein 3D* (id software, 1992), composed and with sound designed by Bobby Prince, who went on to write the music and sound effects for one of the most popular PC games of all time, another FPS game, *Doom* (id software, 1993). *Doom* had a science-fiction-based narrative in which a space marine is sent to Mars where he must fight the demons that have come through the gates of Hell, thanks to a military-industrial conglomerate's experiments. Perhaps learning from *Wolfenstein's* difficulties (in which only one digital sample could play at any one time, so sounds had to be prioritized), sound effects were a key part of the sound of *Doom*, alerting the player to not only the location of enemy demons, but also the type of demon. Prince (2006) described the concept for the sound:

> There were several classes of sounds in *Doom*. One was general active sounds that were not attached to any one demon. These were more or less ambient sounds, but they didn't play until demons close to the player ''woke up'' (usually based upon the player making some noise in the area). Then there were demon active sounds that were attached to individual demons. These sounds let the player know what class of demon was around the corner. Each type of demon had a sight sound that played when the demon ''saw'' the player. There were also attack, hurt and death sounds particular to each type of demon. Another helpful thing about the sound driver was that the volume of sounds depended upon the distance from the player to the source of the sound. This helped keep the overall volume down during noncombat. It also stood to help scare the pants off the player when a demon in a dark niche woke up and immediately screamed his attack sound.

A similar use of sound effects can be heard in the popular 3D stealth game *Thief: The Dark Project* (Eidos, 1998). Sound cues helped to inform the player of nearby enemies, and also helped to inform nearby enemies of the player. Guards usually alerted the player to their presence through sound by walking heavily, singing, or whistling. If the player, playing as thief Garrett, made too much noise, nonplaying characters would be alerted to his presence. The player had to walk softly, muffle footsteps with moss and stay away from hard pavement, grates, or tiled floors. If a guard or other enemy was alerted, the player had to remain still and hide until the enemy gave up searching for him. Other stealth games have followed with similar ideas, such as *Splinter Cell* (Ubisoft, 2002), in which the music ramps up if the player has been discovered. These types of games rely heavily on surround sound technology, helping the player to locate objects and people in a three-dimensional space. They have also elevated the role of sound design and clearly reinforce the idea that playing a game with the sound left switched on is often critical to success.

Myst (Cyan, 1993), an interactive mystery puzzle game, brought a new audience to gaming.[6] People who had not previously been interested in games began to see the allure of a beautifully invented world with difficult puzzles to be solved as part of a narrative. *Myst* was a first-person puzzle in which the player moved about by clicking on the main display and interacting with various objects. Solving puzzles led to new ''Ages,'' or self-contained worlds with further puzzles. One of the attractions of *Myst* was its unusual soundtrack. Rather than rely on standard melodic music (which would have undoubtedly become repetitive in the many hours it took to solve each puzzle), the game used minimal ambient music and sound effects to fill the silence and create an eerie atmosphere. As composer Robyn Miller described,

> Actually, we were weren't going to put music in *Myst*. Imagine *Myst* without music! Rand and I always wanted to create a world. We wanted to put . . . you, the player, in that world, and make you feel as if you were really there. We knew the *Myst* world needed to be as believable as possible and we felt that music was not part of that equation. So, instead of music, we wanted to create an aural mood with sound effects . . . the main focus became making the music as atmospheric or ''environmental'' as possible. . . . For me, as a composer, it was always about creating a more believable virtual world experience for the player.[7]

Minimal musical motifs were heard with ambient effects as the player entered specific locations, or when specific puzzles were solved or devices activated. For instance, in figure 4.1, the musical melody motif near the start of the game (which is repeated throughout this level on various instruments) is played while the player is on the stairway leading up to the telescope platform. Here the theme is on panpipes with a cello-like synthesizer and brief piano flourishes.

Figure 4.1
Myst, "Above Stoneship (Telescope Theme)" (Robyn Miller, Cyan, 1993).

Played at a slow pace, such a minimal approach was beneficial to a game like *Myst*, in which players may be stuck for a significant amount of time at any one stage. Avoiding too "memorable" a theme, or any sense of real "beat," was important to avoiding listener fatigue and to not distract the listener from the activity of puzzle solving. These brief themes would typically loop repetitively until the player changed location, and then would hard cut out to the ambient sounds, illustrating one problem with the rapid jump to Redbook technology: the dynamic MIDI techniques had been abandoned in favor of a return to linear tracks and loops.

With players annoyed by a game's repetitive looping, several methods were explored. One approach to keeping the player interested in and attuned to sound was allowing the player to select the music, such as in *The Sims* (Maxis, 2000; certainly not the first to use this ploy, but by far the most popular). *The Sims* was a spin-off of other popular simulation *God games* ("in which the player oversees the development of an entire city or civilization from a near-deific vantage point," Kline, Dyer-Witherford, and de Peuter 2003, p. 270), including *SimAnt* (Ocean, 1991) and *SimCity* 2000 (Maxis, 1993). *The Sims* was one of the most successful games ever released, with over fifty million games or add-ons sold worldwide.[8] The game was a kind of virtual dollhouse simulation of everyday life of various virtual people, called Sims. The player had to organize their sleep, work, food, and social activities. Kline et al. (2003, pp. 269–293) note how *The Sims* took advantage of the Internet to offer customers the opportunity to add input to the game, through trading items, telling stories, and so on, thus pioneering user-generated content in new ways (beyond modding—that is, modifying using mod packs). In terms of sound, players could select the background music (once a stereo system had been purchased), which allowed for rock, classical, country, or Latin, and included several songs to choose from in each category. The more recent release *The Sims 2* (EA, 2005) even allowed users to include their own MP3s in a subfolder that then allowed the player to play the songs through a radio in the game. In-game music for *Sims 2* was composed by Mark Mothersbaugh of Devo (released on Nettwerk, 2005), with additional songs licensed from popular bands and re-recorded in the original ''Simlish'' language. Although popular artists had been involved in game soundtracks previously (see chapter 6), *The Sims* series perhaps represents the first time these artists went back to the studio to re-record songs in a new language specifically for a game.

HOME CONSOLE AUDIO MATURES

The first 32-bit home consoles, the Panasonic FZ-1 3DO and the Atari Jaguar, were released in 1993. The CD-ROM-based 3DO system was never really popular with gamers, in part because of its high price, although it had truecolor graphics,[9] full motion video, and 16-bit CD-quality audio. Likewise, the Atari Jaguar also had advanced architecture but saw little significant success. There were, however, a few popular titles, most notably a port of *Doom*, and *Tempest 2000* (Atari, 1994). *Tempest 2000*, a remake of an old Atari classic vector-graphic tube shooter,[10] was so popular, in fact, that it was one of the first games to have a separate, commercially sold soundtrack, composed in MOD format by Ian Howe of Imagitec Design. Nevertheless, two other 32-bit consoles would subsequently be released which would have significant mass appeal: the Sega Saturn and the Sony PlayStation.

Sega released their 32-bit Saturn in 1994, a CD-ROM-based machine with eight processors, two of which were reserved for audio. The Saturn Custom Sound Processor (SCSP), manufactured by Yamaha, consisted of a 16-bit DAC and a 32-channel PCM sound generator, capable of CD-quality sample rates. The main drawback to Saturn's sound system was the limited amount of RAM allotted to sound. Because audio samples had to be downloaded raw (decompressed) into the audio memory buffer, this meant that there was a limited amount of space for simultaneous sounds; so the sample rate was often reduced to conserve memory, thereby reducing the fidelity of the sound. The Saturn saw its highest popularity in Japan, where it sold close to six million units—twice the combined sales of the rest of the world. One of the reasons for the ultimate failure of the Saturn in North

America and Europe was the competition with the Sony PlayStation, a system that was cheaper and easier to program for, and thereby had the support of many more games designers.

The Sony PlayStation had begun its life as a CD-ROM add-on component for Nintendo's SNES system. Nintendo had joined forces with Sony to better compete with Sega in the video games market, but the two companies could not agree on the system, and Nintendo eventually signed a contract with Philips. Sony decided to press ahead with its own 32-bit system, the PlayStation (see Kent 2001). Undercutting Nintendo in price, the PlayStation was enormously successful, selling over 85 million units. The PlayStation's CD-ROM drive could also play audio CDs, and, in fact, there were some games in which it was possible to pause the game, and stick in an audio CD to substitute for the games' audio (such as *Twisted Metal 4*, Sony, 1999). Like the Sega Saturn, the PlayStation also offered MIDI support for ease of programming. The sound chip was capable of twenty-four channels of CD-quality sound, and allowed for real-time effects like ADSR changes, looping, reverb, and pitch modulation.

Despite the advances in sound fidelity, in some ways, the PlayStation was a step backward in terms of the way music was presented in games. As on personal computers, Redbook audio meant that there were compressed audio files with more channels and higher-quality sounding instruments used (as well as the ability to use instruments not included in the GM MIDI specs, such as sound effects and vocals), but this was at the cost of dynamic adaptability and interactivity, and, as on PCs, many games went back to rapid fades and hard cuts between tracks, which typically looped continuously in game levels. Not all PlayStation games depended on Redbook audio, however. One notable exception was Nobuo Uematsu's soundtrack for *Final Fantasy VII* (Squaresoft, 1997, figure 4.2), which relied on MIDI from the on-board synth chip.[11] The use of MIDI in the game meant that there could be more room for very dynamic music without having to loop endlessly (the four hours of music was later released on four CDs), and it also freed up some of the CPU time for the extensive 3D graphics in the game. More important was that it allowed for quicker transitions between tracks and a more dynamic score.

After dropping out of their deal with Sony, Nintendo bypassed the 32-bit machines altogether, going straight to a 64-bit release in 1996, the Nintendo 64 (N64). The product's lack of available games in comparison to the Sony PlayStation eventually meant that the N64 would not see the dramatic kind of sales achieved by Sony, although it did reportedly sell thirty million console units.[12] The N64 also well surpassed the PlayStation in technical capabilities in many ways. The main processor controlled the audio and was capable of producing 16-bit stereo sound at a slightly higher sample rate than CD quality (48 MHz). Some games supported surround sound, and this was enhanced further by the third-party add-on release of RumbleFx 3D Sound Amplifier, a device that could help

Figure 4.2
Final Fantasy VII, "Battle music" (Nobuo Uematsu, Squaresoft, 1997).

to mimic surround sound on a stereo system. DSP filters and effects like chorus, panning, and reverb could also be implemented in the internal CPU, or in the software, in real time.

Whereas the PlayStation had been reliant to a large extent on CD audio, Nintendo stuck with a General MIDI-based system. There were several custom sound programs for Nintendo developers to compose music, one of which was known as MusyX. MusyX offered its own programming language, SMaL, which could influence the way a sample was played back. Unlike DSP-based parameters, the waveforms in SMaL could be changed in relation to MIDI controller values. MusyX offered wavetable synthesis with samples the composer could input, in a MOD-like format. Similar to iMUSE, markers could be set in sequences and multiple sequences could be played simultaneously to allow for cross-fading, or layering of sounds with a mapping program. Not all games used the MusyX system, however, and the degrees of dynamic audio in games varied considerably. There were, however, many examples of popular games that were still considerably dynamic. *Banjo and Kazooie* (Rare, 1998, music by Grant Kirkhope), for instance, had a dynamic MIDI-based score, which changed instruments in the track as the player moved about various locales. *The Legend of Zelda: Ocarina of Time* (Nintendo, 1998, music by Koji Kondo) used a variety of dynamic approaches, discussed further in chapter 8. For instance, when the player's character encountered a threatening enemy, a subtle cross-fade occurred between the game-play music and the threat-music. Although critics may have disparaged the MIDI audio, in terms of advancing a dynamic approach to game sound, Nintendo was leagues ahead of its competitors in the console domain.[13]

In the face of Sony's and Nintendo's success, by the late 1990s Sega was in trouble: They had achieved unexpected success with the Genesis, but their systems since that time had floundered. They introduced the first 128-bit console, the Sega Dreamcast, in 1998, with a disappointing consumer response, despite impressive capabilities (see Kent 2001, pp. 563ff). The sound had two processors, and, unlike the Nintendo 64, samples did not have to be decompressed, improving real-time audio fidelity. More important was that the dedicated audio processors had their own memory, so that sound quality was not compromised by other aspects of the game. The Dreamcast reportedly sold only about six million units until it was discontinued in 2002, and with the release of PlayStation 2, many of Sega's top designers jumped ship, dooming Sega in the console realm.[14] The 2000 follow-up to the immensely popular PlayStation had ensured that fans of the original system would be suitably impressed with the new machine. With the ability to play DVD movies and the option of add-ons for modem and hard drive, the PlayStation 2 took an important step toward becoming a ''home entertainment center.'' Its games were stored on DVDs capable of holding 5.7 gigabytes, and it fully supported the multichannel surround sound standards AC-3, DTS, and Dolby Digital, offering up to eight separate speaker channels (see box 4.1, ''Dolby Formats''). The sound-processing unit was capable of 16-bit audio with a maximum sample rate of 48 kHz—better than CD audio—and had an additional 48 MIDI channels. However, limitations still meant that sound quality had to be compressed to save space, with the result being that cinematic scenes typically had high-quality full surround sound, but when the player began to use other resources requiring real-time processing, the music and sound would take second place and often drop down to two-channel stereo. This was improved by the PlayStation 3 (PS3), Sony's immensely popular follow-up to its PlayStation consoles, released in 2006. The PS3 is able to run up to 512 channels and apply different layers or DSP filters in real time. It is also capable of streaming audio in 7.1 at 96kBs, though reduced channel formats (5.1, for instance) have higher bit-rate streams, which is particularly important for Sony's online service for multiplayer games. Nevertheless, the processor, known as the Cell, also handles graphics and other game functions, such as AI, meaning that audio must still compete for shared system memory and CPU processing.

Box 4.1
Dolby and Surround Sound Formats

DOLBY DIGITAL (also known as Dolby AC-3) is the current standard for home theater surround sound, and also present in many movie theaters—the 5.1 surround system. Dolby Digital supports up to 5.1 channels, and therefore can also refer to mono (1.0), stereo, or Pro-Logic (2.0), or five channel audio (5.0). Specifically surround formats are referred to as ''Dolby Digital 5.1.'' Dolby Digital 5.1 is the standard for DVD-video and HDTV (high-definition television), used by pay-per-view and digital TV channels. It provides up to five independent channels (left, right, center, surround left, surround right), plus an optional sixth (LFE, or low frequency effects), typically handled by the subwoofer. Since the sixth channel is not full frequency (it only handles deep bass, 3 Hz to 120 Hz), it is referred to as ''*.1*.''

DTS DIGITAL SURROUND is a competing format to Dolby Digital, a 5.1 surround system available in movie theaters, and optional on some DVD movies for home theater. DTS is not a standard soundtrack for DVD or HDTV, but offers higher data rates than Dolby Digital (that is, it uses less compression), resulting in more data demands, and hence is not as useful for games as Dolby Digital.

DOLBY SURROUND PRO-LOGIC is the standard for Hi-Fi VHS or analog television. It is encoded in stereo (or stereo analog). Pro Logic II can be found on most new receivers compatible with all stereo and matrix-encoded surround software. Pro Logic II is an update of the original Pro Logic and can be used to freshen up Dolby Surround-encoded videotapes and DVDs. It also works wonders with stereo CDs and radio broadcasts.

DOLBY DIGITAL PLUS provides up to 7.1 channels of audio, comparable to DTS-HD high resolution audio, and is ''better than DVD'' quality sound (in that it has a higher bit rate and new coding efficiencies).

DOLBY TRUE HD also provides up to 7.1 channels, lossless audio; is equivalent to DTS-HD Master Audio, and was developed for high-definition TV and Blu-Ray disc players.

DOLBY DIGITAL EX is another ''extended surround'' format for home theater, equivalent to THX Surround EX. To a regular 5.1 configuration, it adds at least one (typically two) back channels, with speaker(s) behind the audience, allowing sound effects behind the audience, resulting in 7.1 surround.

Nintendo had responded to Sony's PS2 with their Game Cube in 2001, but the Game Cube failed to recapture the market that Nintendo was looking for. The competition was fierce, and the game titles for the Nintendo were more family-oriented, typically appealing to a younger market. The Nintendo Wii, released in 2006, however, distinguished itself immensely from competitors through its unique control system, a handheld remote with motion sensors. Designed to bring new gamers into the world of video games, Wii advertising focuses on the populations typically ignored by many games manufacturers—the very young, the elderly, and families. Although the sound capabilities of Wii are less than its competitors,[15] Nintendo instead created a novel idea in the form of a tiny speaker in the controller. Various in-game actions result in sound emanating from the device, creating an even more dynamic, immersive audio experience. The sound effects of the bow and arrow in *The Legend of Zelda: Twilight Princess* (Nintendo, 2006), for instance, give the impression of an arrow being shot from the player toward the screen.

The success of home consoles for a time was viewed as a threat to the PC market; after all, consoles were becoming full multimedia devices, and many people who had used their PCs for little more than playing games and surfing the Internet could do these through a much cheaper console. It was no surprise, then, that rumors of a games console from Microsoft began around 1999, and a system was officially announced in 2000. Microsoft's entry into the console business the following year, the Xbox, was built around a Pentium III processor with an 8 GB hard drive for music, graphics, and saved-game information. Games were supplied on 5.7 GB DVD discs, and the Xbox was also capable of playing DVD movies and audio CDs, making it a full multimedia device. The Xbox featured its own audio processor, known as SoundStorm, which would support Microsoft's DirectX protocol.[16] Xbox Live was an important part of the console's success, an online multiplayer subscription service which allowed for episodic content (downloadable new content for games), and access to Xbox Live Arcade, which included classic arcade games such as old Genesis and PlayStation games. The upgraded Xbox 360 (2006) requires that developers support at least Dolby Digital Live 5.1, as well as allow gamers to use their own playlists (from a port allowing a plug-in of an MP3 player).

OTHER PLATFORMS: RHYTHM-ACTION, HANDHELDS, AND ONLINE GAMES

In addition to the arcades and home market, video games have also been developed for other platforms, including specially controlled rhythm-action games, handheld consoles, mobile telephone games, and games designed for the Internet.

A brief overview of this alternate gaming world will round out the history presented thus far.

Rhythm-Action Games

Rhythm-action games are video games in which the player must respond in some way to the rhythm or melody being presented, either through repeating the same melody or rhythm by pressing buttons (with hands or feet), or kinetically responding in some other way to the rhythm, often using specially designed controllers. Rhythm-action games to some extent have been incorporated into some other types of video games—in several of the *Legend of Zelda* games, for instance, the player is required to memorize short melodies and play them back (*Ocarina of Time*), or have their character ''dance'' to various rhythms (*Oracle of Ages*, Nintendo, 2001). In chapter 3 I discussed one such game, *Toejam and Earl: Panic on Funkotron* for the Sega Genesis, which included a percussion memory game. Although these games integrated elements of rhythm-action games, these elements were just one small component of the game, and so they are not discussed further here.

Atari was one of the first companies to release an electronic rhythm-action game, with their *Touch Me*, designed by Ralph Baer and released as both a coin-op in 1974 and handheld in 1978. *Touch Me*'s original 1974 flyer discussed the use of sound: '''Beep' or 'Bleep'? *Touch Me* challenges the player to remember the sequence of sight and sound, and correctly repeat the pattern ... When the sound occurs, the corresponding button lights to give a visual clue ... Did you correctly repeat those sounds when the button lit?''[17] There were four buttons and four tones on the machine whose patterns would get progressively more difficult. *Touch Me* did not catch on, however, until it was redesigned by Milton Bradley and released as *Simon* in 1977. Similar to *Touch Me*, each button on *Simon* had a corresponding sound and color, and increasingly difficult melodies had to be memorized. A wave of copy-cat games followed, including *Super Simon* (a two-player version by Milton Bradley, in 1980), Parker Brothers' *Merlin* (1978), the Tiger *Copy Cat* (1979), and Castle's *Einstein* (1979). More complex was the *Logix T.E.A.M.M.A.T.E* (Parker Bros., 1980), a tabletop console, which had a speaker and a primitive sequencer that allowed the user to create his or her own melodies, described in a 1980 *Consumer Reports* magazine: ''By way of introduction to this musical entertainment, the toy provides a preprogrammed version of 'Oh, Susanna.' It didn't really encourage us to proceed, since it was rhythmically faulty. Persevering, though, we found music-writing rather a challenge—at first. But we became frustrated when our laborious compositions were erased when the toy was turned off'' (p. 655).

The popularity of rhythm-action games continued into the home console era. Nintendo introduced a ''Power Pad'' to the NES in 1988, a floor mat with control buttons built in for the feet, whereupon games like *Dance Aerobics* (Bandai, 1989) would become forerunners to today's popular arcade rhythm-action games. *PaRappa the Rapper* (SCEI, 1996), released on PlayStation, incorporated kinetic elements into a storyline: PaRappa was trying to win the love of a girl, Sunny, and had to practice rapping to the beat of various drummers to improve his skill. A button was to be hit in the correct order with correct timing to keep the beat, significantly relying on memorization, much like *Simon*. Handheld consoles have also spawned their share of rhythm-action hits, including *Elite Beat Agents* (for the Nintendo DS, 2006) and *Electroplankton* (Nintendo, 2006) in which elements must be tapped or spun with the stylus.

There are often add-on user interfaces or controllers for rhythm-action games. *Donkey Konga* (Nintendo, 2003) included ''DK Bongo'' drums, *Guitar Hero* (RedOctane, 2005) a guitar-shaped controller, and so on. *Beat Mania* (Konami, 1997) included a five-key controller and accompanying turntable that had to be manipulated according to on-screen instructions during a series of increasingly difficult songs. Dancing games, such as Andamiro's *Pump It Up* (1998), Roxor Games' *In The Groove* (2004, which has since been bought out by Konami), *Britney's Dance Beat* (THQ, 2002), and of course *Dance Dance Revolution* (Konami, 1999) are probably the best known of the rhythm-action games, spawning considerable fan communities around the world, and perhaps reviving interest in coin-op consoles (see, e.g., Demers 2006). Other music-based games, including remixing games such as *FreQuency* (Harmonix, 2001) and singing games like *SingStar* (Sony, 2004), have also more recently put the player in a partial role of musician (composer as well as performer; see chapter 6).

Handheld Consoles

Handheld video games (smaller, battery-powered portable consoles with built-in screens) became popular alongside other games consoles and the arcades. Mattel had manufactured handheld LED-based games such as *Missile Attack*, *Baseball*, and others as early as 1976, and as shown above, Atari had released a handheld version of *Touch Me* in 1978. Milton Bradley released the Microvision in 1979—a more advanced type of console, with interchangeable cartridges of games such as *Bowling* and *Pinball*. Early *tabletop* games—small, one-game arcade ports—were also popular in schoolyards in the early 1980s, though these were generally without sound, or very limited in sound capabilities. It is interesting that a 1980 *Consumer Reports* report on games mentions sound in a few cases, indicating that sound played an important role even this early on, although its effectiveness was questionable. The description of Vanity Fair's *Computer Matician 3010* tabletop model, for instance, which asked a series of math questions, included sound: ''The correct answer elicits a high-pitched, repeating beep and the word 'right' flashing at the top of the screen. An incorrect answer is punished by a penetrating low buzz and the word 'wrong' flashing at the bottom of the screen . . . most of the children were unnerved and annoyed by the strident beeping and buzzing'' (*Consumer Reports*, Nov. 1980, p. 654).

Sound improved as the developers of home consoles released handheld versions of their home games to fans. The Atari Lynx (1989) was the first color handheld game, and had four-channel 8-bit sound. NEC released Turbografx Express a year later, which was basically a portable version of the Turbografx16. Both of these models, however, failed to generate any real consumer excitement. Sega's Game Gear was released in 1991, using the same primitive sound chip that was found in the PCjr. The Game Gear was fairly successful, with about 250 games developed for the system, as was Sega's 1995 Nomad, a handheld version of the Sega Genesis.[18] Other companies had some moderate success with handhelds, such as Neo Geo's Pocket and Pocket Color of 1998–99, a 16-bit handheld version of the Neo Geo home console, with a six-channel PSG chip. Without a doubt, however, the uncontested leader of handhelds was the Nintendo Game Boy. Nintendo had released handhelds previously with its Game and Watch system of 1980, but it was Nintendo's Game Boy which would capture the hearts of players, and through its various guises see over 100 million units sold (Wing n.d.).[19] The original black-and-white Game Boy received its first update in 1996 when Game Boy Pocket was released—essentially a more streamlined version of the original but with no other major changes. The second update came with Game Boy Color two years later. Though this model included improved graphics, the sound remained the same for these three models; a three-plus-one channel stereo sound PSG, the fourth channel offering 4-bit noise, with a simple envelope generator. The Game Boy Advance, released in 2001, was the first to upgrade the sound, by adding two 8-bit DACs to the original configuration.[20] Not only the hardware but the software was improved: the Advance BIOS contained many sound-related functions for converting MIDI to Game Boy data,[21] although it is worth noting that Nintendo did not even bother to release information about sound in their press release specifications, suggesting that sound was not considered as relevant on handheld games as it was on home consoles.[22]

The Game Boy had a very distinct sound, which is perhaps one of the reasons why it has been so popular with chiptune musicians and the instrument on a remix album of Beck songs, *The GameBoy Variations* (*Hell Yes*) (Interscope, 2005). On the Game Boy system, the first two channels were square waves, which were commonly used for chords or melody.[23] Channel 3 was a programmable (variable) wave channel with thirty-two 4-bit programmable samples, though it was commonly used for bass lines. Channel 4 was a white noise generator, most often used for percussion sounds, which typically sounded like a bit of scratchy static rather than drums, as shown in figure 4.3.

More recent developments in handheld technology have led to handheld consoles becoming full multimedia devices, capable of playing movies and MP3

Figure 4.3
Aladdin (Seiko Kobuchi and Chihiro Arisaka, Capcom, 2004) for the GameBoy Advance, with fairly typical usage of GameBoy channels.

audio, such as the PlayStation Portable (PSP), with 3D multichannel sound and integrated stereo speakers. The Nintendo DS has even become a musical instrument in its own right, with tracker, MIDI, and sampling software available. It has a built-in microphone, stereo speakers with virtual surround, sixteen hardware sound channels, and two built-in samplers. There are also some interesting musical add-ons for the Nintendo DS, including an X/Y MIDI controller resembling Korg's "Kaoss Pad," MIDI keyboard controller, and wireless DS synthesizer. Overall, however, handheld game audio has lagged behind their console cousins because the games are typically played in a public domain, and as such audio plays a secondary role to graphics (see below).

Mobile Phone Games

The history of mobile phone games (distinguished from handheld gaming in that games are not the primary intended use of the machines) in many ways echoes the history of consoles and handhelds. The earliest mobile phone game, *Snake*, written by Taneli Armanto on the Nokia 6110 in 1997, was black-and-white and lacked sound. The phone also featured basic *Memory* and *Logic* games. These games were originally installed on phones by the phone companies, but are now commonly purchased through mobile networks or through Internet download. As well, mobile phones are increasingly developing toward small computers, known as smartphones, allowing for more complex games, such as Nokia's N-Gage game deck with changeable cartridges and Bluetooth multiplayer gaming for small groups. One of the most important elements of mobile gaming has been network capabilities. Even many early games had multiplayer modes and simple games like checkers could be played with live competitors.

Mobile gaming is rapidly becoming an important part of the games industry, with publishers like Jamdat joining forces with Electronic Arts. At the time of writing, Jamdat leads the mobile gaming world, but Namco is in second place, having re-released its classic arcade games like *Pac-Man* and *Galaga* to the mobile community. Vivendi Games is also plundering old games for material, releasing a series of Sierra back-catalog games in 2007, including *The Incredible Machine* (1992) and a *Leisure Suit Larry* game. It is likely that the popular games of the early consoles and arcades are all to be re-released in the coming years in mobile form, although those games that used cover songs without licensing will likely see their music change.

Mobile games today are comparable to 16-bit home consoles in terms of graphics quality, but audio remains very difficult to produce for a variety of reasons—not least of which is the lack of standards. Many of the same difficulties and constraints that plagued early computer games are now seen in the mobile game audio world. Audio resources on phones can be as small as 100 kB, and as such, many of the same techniques of the 8-bit era have been reproduced to save space. As composer and self-confessed ''annoying ring tone guy'' Peter Drescher elaborates, ''The most important trick is to be as ruthlessly efficient as possible. You want to squeeze every last drop of variation out of each and every byte of audio data at your disposal. Repetition is the enemy, compression is your ally, and creative use of limited resources is your battle cry. If you're making sound effects, this means low-resolution, highly compressed samples used in multiple ways. If you're writing music, it means MIDI'' (Drescher 2006).

Mobile audio also has different requirements from other games owing to the casual nature of the games—people playing mobile games are, at present, primarily using their phone to play games while they ''kill time,'' and games are secondary to the other functions of these machines. Nokia's guide to sound for phone games, for instance, warns, ''The game should be playable without the sounds. Allow silent starting of games. If intro music is implemented, there must be a way to switch it off. Prompt for sound settings at the start of the game. . . . Do not include loud or high-pitched sounds, have sounds on as the default, [or] use sounds that are similar to ring tones and alert tones'' (Nokia Corp. 2005). On the other hand, Nokia reminds developers, ''From a cognitive point of view, the auditory channel is an excellent way to relay information and give feedback to the player. While the background music can be used to enhance the emotions and mood of the game, a specific sound effect can be used to communicate meaningful game information'' (ibid.).

Real-time network play on mobile phones has become increasingly popular, with the introduction of games like *Bejeweled Multiplayer* (Jamdat, 2006). Some games can also interface with their PC counterparts, such as *Ragnarok Mobile Mage* (Skyzone Mobile, 2006), so that players can continue their game after they have left their house.

Online Games

The history of online gaming can be divided into roughly two categories:[24] casual games (easy-to-learn games with mass appeal like *Sudoku*, *Tetris*, and so on), which are usually one-player (although there are competitive casual games, including those packaged with MSN Messenger, for instance); and multiplayer worlds, including those called MUDs (multiuser domains or dungeons), MMORPGs (massively multiplayer online role playing games), MMORTS (massively multiplayer online real-time strategy), and MMOFPS (massively multiplayer online first-person shooters). There are other online games that do not really fit into these categories, such as the online version of *Sims*; however, for the sake of brevity I will refer to all of these multiplayer games as massively multiplayer online games, or MMOs.

Casual games have been around for a long time on the Internet, from HTML-based quiz games to more elaborate Java or Flash-based games.[25] The present preference for developing web-based games is Macromedia Flash, and so I will focus on Flash for the remainder of this discussion. Flash was in fact created by a game developer, Jonathan Gay, who had developed PC games in BASIC but initially enjoyed success with *Airborne* (Silicon Beach Software, 1984), one of the first Macintosh games to use digital music (*Ride of the Valkyries*, during the opening sequence). With this history, Flash was developed with the notion of creating a smooth animation and sound program in one package. Since becoming part of Macromedia, Flash has become far more complex and developed its own scripting language, ActionScript, which allows for more complex programs.[26]

With Flash, developers can create games and animations with synchronized soundtracks, although synchronization between sounds is difficult. Using a timeline-based approach, sounds that begin simultaneously will stay in sync, and so tracks can be kept separate. Synchronization between layers of sounds can be unreliable, however, particularly on slower computers, and synchronizing audio to visuals can also be difficult. Streaming sound over the web can mean that visual frames will sometimes be skipped in order to remain in audio sync.

One of the more interesting uses of sound in contemporary casual games comes from *Warbears: Mission 2* (composer Filipe Ferreira, 2006).[27] The game is basically a Flash based point-and-click puzzle that takes place on a single screen. There are four characters (bears) that the player controls to work together to solve a puzzle. The music is essentially one long loop,[28] but different segments of gameplay elicit different melodic elements and sound effects, which add a layer overtop of this loop. When the player clicks on the bear Ryoh to cut down a tree with his sword, for instance, a panpipe-like melody plays overtop of the main loop (figure 4.4). When the player clicks on the bear Lucas to push switches to control the alarm in the house, a new bass line is added.

The other significant form of online gaming is massively multiplayer online games (MMOs), which have become a tremendous part of the game industry. The history of MMOs goes back to at least 1969, when a two-person version of *Spacewar!* was written for PLATO, an early file-sharing system developed at

Figure 4.4
Warbears: Mission 2: showing three variations which are added over the main percussive loop: the panpipe melody, a guitar stutter, and timpani-like sound (Filipe Ferreira, 2006).

the University of Illinois. Over the next decade, many MMOG games would be developed for the system, including a *Star Trek*–based game called *Empire*, a flight simulator called *Airfight*, and several *Dungeons and Dragons* games. Developments that would allow users to kill one another (*Mazewar*) and communicate with one another (*Dungen*) soon followed. PLATO went through a series of advancements, the most relevant to this discussion being the PLATO IV computers from 1972, which had a ''Gooch Box,'' a four-voice synthesis chip. Later, a Gooch Cybernetic Synthesizer, a sixteen-channel chip, was added.

Other universities developed online games, and by the late 1970s the creation of multiuser dungeons, or MUDS, formed, beginning with MUD1, in 1978, a real-time multiplayer game with many of the characteristics of MMOGs today, most notably the idea of character development through experience (see Cox 1994). Commercial response was soon to follow, with games like *MegaWars* (1983) and *Islands of Kesmai* (1984) on CompuServe. LucasArts was quick to join in, with a release of *Habitat* for the Commodore 64 Q-Link (1985). *Habitat* had color graphics and some sound effects, such as dial tones when the player picked up the virtual phone. With the rise of the World Wide Web in the 1990s came a new series of online games that charged subscription fees and had bigger budgets, but music remained relatively poor in quality in comparison to offline games as a result of music's file size and the nature of downloading through modem. Three games sometimes referred to as the ''Big Three'' dominated the 1990s: *Ultima Online*, trading on the success of the offline *Ultima* games, was released in 1997 (Electronic Arts); *EverQuest*, launched in March 1999 (by Verant Interactive, later acquired by Sony), and Turbine's *Asheron's Call* (1999). More recently, Linden Lab opened up its innovative ''3D virtual world'' called *Second Life* to the public (2003), a game that quickly became the most popular non-combat-based-MMO. I take up compositional approaches to MMOs in chapter 8.

CONCLUSION

The next generation (''next-gen'') games systems (referring to the most recent—at time of writing—generation of consoles, including Xbox 360, PlayStation 3, and Nintendo Wii) are much closer in terms of capabilities to home computers than their predecessors, reducing further the line between home entertainment systems and home computers. Arcade games have tried to lure players with ever more sophisticated playing consoles that include rumble effects and built-in full surround sound.[29]

Says producer and audio supervisor Mike Verrette of Ironlore Entertainment, ''The PC used to be capable of so much more than the original Nintendo and the other early consoles, but that gap is smaller now'' (Lehrman 2007, p. 22).

Some might even say that next-gen games represent an end to the technological audio constraints of the past, as certainly, many of the technological problems have been solved. However, constraints do remain, and there are still many difficulties to overcome in the production of game audio, including problems of budget, time, and increasing expectations placed on game audio by consumers and developers. Technological problems also remain, particularly in that the development of software has failed to keep up with the upgrades in hardware. Most sequencing software is still loop-based and linear, for instance, and there are few affordable cross-platform *middleware* audio solutions.[30] With hardware companies each competing to introduce their technologies as standards, resulting in many different requirements for each console system, cross-platform middleware software has been slow to develop, or tends to remain proprietary (such as Sony's SCREAM engine). In fact, developing a title for all three consoles can create significant problems since tools and compression technology, for instance, can vary from one system to another, and solutions created for one console often do not work on another (see Lehrman 2007, p. 22).

Another remaining problem in consoles is audio streaming. In streamed audio, music is sampled into long (typically stereo) files stored on the CD/DVD and played directly in real time in the game, through a RAM buffer. Although this often cannot work for downloadable games (since the files are too large), streaming audio off a disc allows for more real-time processing in console games. The problem, however, is that speeds are still relatively slow and demands on the processor means audio must compete with other game aspects, such as graphics, reducing channels or otherwise compressing audio requirements. Typically, sound is allowed to occupy only 5 to 10 percent of system resources. Jason Booth, designer at Harmonix, elaborates:

> On PlayStation 2, you have two megs of RAM for audio. All of the sound for wherever you are in the game has to reside in that memory. Sometimes you have to leave room in there for a streaming buffer, but you also have to worry about the speed of the disc spin. Data on the outside of the disc reads faster than data on the inside, so you have to take that into account—data for loading a new level might be on the inside, while data used during run-time is placed on the outside. A common technique is to load in the sound effects ahead of time and stream the music. But of course, you have to be okay with the rest of the team with that: They can't be loading textures at the same time you're streaming music. (Cited in Lehrman 2007, p. 22)

Streaming technologies can be expected to advance in the coming years, although they are already greatly improved over the previous generation of consoles.

Particularly important to the development of games today is the convergence of the various platforms of media and technology: I can play games and check my email on my phone, connecting it to my computer to continue playing. I can connect to other gamers online through my Xbox, my computer, or my phone. I can watch movies on my PlayStation Portable, compose music on my Nintendo DS and upload it to friends online through my PC or email it to a friend so that they can use it as a ringtone on their phone. As technologies for portable systems continue to develop at a frenzied pace, this collapse of media platforms into multiuse devices will become even more important, and we can expect games to play an increasingly important role for portable devices. In many ways, games have driven and will continue to drive these technologies, particularly in the area of sound.

As discussed in the previous chapter, the rapid development of technologies has depended to some extent on the creation of and agreement on standards. In the case of audio, groups such as the Interactive Audio Special Interest Group (IAsig) have been working toward these means by creating and recommending such standards, although the industries involved still work very much on a proprietary level, which arguably slows progress. One interesting future development for games may in fact be the convergence of different platforms altogether, as Silicon Knights' Denis Dyack suggests, the result of which will mean more time put toward the development of games, and less time wasted porting games between systems:

> I think we're moving towards a homogeneous platform whether people like it or not. At the end of the day, I think it's in everyone's best interest that there be one hardware console, whether it be Sony, Microsoft, Nintendo or whether all three of them got together and said, "Ok we're going to agree upon a standard for everyone to make." In the movie industry it helped tremendously because as a content creator, all we want to do is make games and entertain people. Don't get me wrong, I love the hardware platforms, like the Sony platform and I think the Wii's got some really unique things and Microsoft's platform we obviously love a lot. However, we'd rather spend time making the games than worrying about the hardware. And if everyone had the same hardware and when you made a game you knew you got 100% penetration because anyone who plays this game had to buy this hardware platform just like a DVD or whatever standard media format's going to be. I think that would ultimately be much better for gamers. If the value of that technology continues to diminish, and it's becoming more and more expensive to manufacture and research & develop these, eventually there's going to come a breaking point where everyone goes, "You know what, let's stop spending all this money on R&D into this hardware where no one really cares about it as much as they used to because of the value of it. Everyone's got one, so why don't we just spend our money on making games because all the money is in the software anyway?" So I think it's going to get to the point where the value of the hardware, like a PS3 or an Xbox 360 or a Nintendo Wii or whatever in the future, is going to be very low. (Cited in Brightman 2007)

The many different platforms and forms of video games have all traveled similar trajectories in terms of audio production, from no sound, to basic linear

MIDI in one channel, to multichannel dynamic audio in surround, with complicated approaches to sound playback. But the history of games, as shown, is itself nonlinear, influenced by industry, technology, and social needs, knowledge, and desires. As Jean-Louis Comolli wrote of film, it is a ‘‘stratified history, that is, a history characterized by discontinuous temporality’’ (cited in Belton 1992, p. 159). Technological developments throughout its history have influenced not only the ways in which game audio has been produced, but also the ways in which it has been received by its audience. As in early film sound, ‘‘The expectations of audiences—and what they perceived as ‘realistic’—were changing year by year’’ (Belton 1999, pp. 234–235). Moreover, as in film, ‘‘each new technological development (sound, panchromatic stock, color) points out to viewers just how ‘unrealistic’ the previous image was and also reminds them that the present image, even though more realistic, will also be superseded in the future—thus constantly sustaining the state of disavowal’’ (Manovich 2001, p. 186). The development of game sound has represented an ever-increasing drive toward greater fidelity and higher realism, a subject I take up in chapter 7. However, the drive to realism is not the only determinant in the development of technology, as shown, and it is now necessary to examine in further detail the industrial, commercial aspects of the industry and the processes by which game audio is created.