# 1

# The Myth of Perfect Fidelity

If you listen closely to the "The Colour of Spring," the opening song of English musician Mark Hollis's self-titled solo album, you first hear nothing.[1] Then, slowly, tape noise fades in. It softly rustles and crackles, before, after eighteen seconds, the first piano chord is struck. The noise remains audible throughout the song, swaying like the sound of gentle summer breeze or a creek not far off. Although sometimes it is almost covered over by the rest of the music, the noise can always be heard—especially when listening on headphones. One could argue that this more or less continuous layer of noise sonically "frames" the other sounds, as it were, thus remaining separate from the actual music. But that would be too easy: given its prominence in the song's first twenty seconds, it is clear that this is more than just background noise. In the late 1990s, when "The Colour of Spring" was recorded, this amount of tape-noise had already been perfectly avoidable for decades. Furthermore, the noise does not become disruptive at any point, the other sounds being rich, subtle, well-rounded, and refined. One must conclude, then, that its presence is a deliberate musical gesture and integral part of the song's sonic template. As part of the music itself, the noise contributes to the way in which the music resonates with listeners.

This use of tape noise in "The Colour of Spring" is just one, especially clear example of how composers, musicians, and recording engineers have creatively explored the unavoidable sonic "side effects" of technological sound reproduction over its one-hundred-forty-year history. From the earliest days of sound recording, people adjusted vocal and instrumental techniques to exploit or emphasize the sonic particularities of technologies, or they actively sought out ways in which certain technical devices change the sound.[2] Whether out of necessity or aesthetic curiosity, they explored how microphones color a singing voice, how sticking a pencil in an amplifier causes an explosion

[1] "The Colour of Spring," track 1 on Mark Hollis, *Mark Hollis*, Polydor 1998, compact disc.

[2] Regarding the first category, Mark Katz gives several poignant examples of ways in which musicians adjusted their playing to match the requirements of recording equipment: from a singer "literally stick[ing] her head inside the horn to ensure that her pianissimo would be heard, but then [ . . . ] quickly withdraw for her fortissimo," via string instruments being replaced by brass "for they could play louder and their sound was more easily directed toward the recording horn," to musicians singing or playing with extra restraint to ensure a well-balanced recording. Katz, *Capturing Sound*, 44–45.

*The Logic of Filtering*. Melle Jan Kromhout, Oxford University Press (2021). © Oxford University Press.
DOI: 10.1093/oso/9780190070137.003.0002

of harmonics, and how slightly oversaturated tape creates a much warmer sound, to give just three examples. Still, such examples notwithstanding, the history of sound technology can just as easily be told as a story of the fight *against* noise. Indeed, from early wax cylinders to the most advanced digital technology, strategies to prevent, reduce, or eliminate the influence of noise have been a major concern among inventors, developers, and engineers.

To confront this apparent contradiction between technological ideals and musical attitudes toward noise, this chapter offers a comprehensive history of the technological struggle with and against noise and distortion. It provides neither an exhaustive history of sound recording nor an analysis of the history of noise as such. Rather, it explores how the concepts of fidelity, mediation, and sound definition run through the history of sound media and define attitudes toward noise.

This history roughly divides into three phases. The first corresponds to the era of disc recording, which starts with the development of acoustical recording up to World War I. At this time, the prevention and reduction of noise was primarily seen as a question of improving and refining the technologies themselves. This phase continues with the introduction of electrical recording in the 1920s, which allowed for the application of noise reduction technologies that had already been developed for other communication media such as the telephone, telegraph, and radio. The second phase began in the 1940s and 1950s. Although magnetic tape recording dramatically improved the quality and flexibility of recordings in this period, it also introduced "tape hiss," which necessitated the development of more sophisticated noise reduction strategies. In the third phase, which began in the 1970s and 1980s and runs up to the present, the advance of digital sound equipment seemingly achieved the ideal—so long sought-after—of completely noiseless sound (re)production. As we shall see, however, even digital recording methods have physical limitations, which introduce specific types of noise and distortion.

Across these three phases, then, approaches toward noise and distortion developed significantly. From early attempts to improve the recording and reproduction apparatuses in themselves during the formative years of sound recording, they have moved through the introduction of increasingly sophisticated methods of reducing the influence of noise between the 1920s and 1960s (culminating in the systems of Ray Dolby in the early 1960s), and toward the unprecedented but by no means unproblematic levels of noise reduction, prevention, and removal that mark the digital age. The myth of perfect fidelity emerges through this continuous back-and-forth between new technologies, new forms of noise, and new ways of dealing with noise. According to this myth, the history of sound recording technology is a slow but steady progression toward noiseless, fully transparent recordings. My genealogy, however, shows that this desire for absolute transparency and perfect fidelity is a limited and ultimately unproductive way of understanding recorded sound and music.

## Mechanical Improvements: Disc Recording

The revolutionary machine unveiled by Thomas Edison in 1877 proved decisively that sound reproduction is technologically possible. Nevertheless, as music critic Roland Gelatt remarked in 1954, it was still "an instrument of crude design and dubious utility."[3] Yes, the phonograph could capture and (re)produce physical sounds, but by today's standards, it was not all that good at it. Sound waves were captured by a large horn and transduced into mechanical vibrations via a diaphragm. Moving vertically up and down, the stylus—a needle directly attached to the diaphragm—etched the vibrations onto a piece of tinfoil wrapped around a metal cylinder. During playback, a different diaphragm turned the vertical indentations back into sound, amplified by a playback horn.[4] In early models, a manual crank was used to rotate the cylinder, meaning that the operator needed a very steady hand so as to keep the recording speed more or less constant. It was a noisy affair. The stylus scratched on the tinfoil, the hand crank squeaked, and the cylinder buzzed. Most efforts at improving the sound quality of early phonography, therefore, aimed to reduce this flurry of machinic noises, whether by experimenting with different recording materials, fine-tuning mechanical parts, or improving the design and functionality of inscription surfaces, styluses, and horns.

From about 1880, Alexander Graham Bell—who patented a working prototype of the telephone in 1876—sought to improve the phonograph, together with his cousin Chichester Alexander Bell and colleague Charles Sumner Tainter. Presented in 1887, their "graphophone" realized a first significant step in sound quality. It substituted Edison's metal cylinder wrapped in tinfoil with a cardboard cylinder coated in wax and made significant improvements

[3] Roland Gelatt, *The Fabulous Phonograph: From Tin Foil to High Fidelity* (New York: J.B. Lippincott Company, 1954), 26. Andre Millard describes how audiences at early demonstrations of the phonograph "had to pay close attention to discern the faint noises coming from the vibrating diaphragm." Andre Millard, *America on Record: A History of Recorded Sound*, Second Edition (New York: Cambridge University Press, 2005), 26. Millard's anecdote supports Sterne's argument that, at a time "when sound-reproduction technologies barely worked, they needed human assistance to stitch together the apparent gaps in the ability to make recognizable sounds." At the earliest stage of sound recording, listeners had to put some trust in the machine to *believe* it could indeed do what was claimed it could and be able to classify the faint sounds that emerged from the horn as bona fide sound reproductions. Sterne, *Audible Past*, 246.

[4] Gelatt, *Fabulous Phonograph*, 20–21.

to the stylus.[5] Edison, in turn, responded with an improved version of the phonograph, while Emile Berliner developed the competing "gramophone." Using discs instead of cylinders, the gramophone applies a "lateral cut" that etches sound waves into the recording groove in a horizontal instead of a vertical way. Because the sound quality of Berliner's technology did initially not match that of cylinders, the latter "outproduced and outsold discs through 1911."[6] After the introduction of shellac discs in 1897 and more user-friendly gramophone players in 1901, however, even Edison's 1912 "Diamond Disc" (boasting a diamond-tipped stylus and new recording material) could not prevent the disc format from taking over the market.

Alongside strictly technological improvements, recording engineers also employed strategies to optimize the recording process itself and catch as much fleeting sound as physically possible. They experimented, as Emily Thompson describes, with the relative positioning of musicians in a room, "the selection of different sizes and shapes of horns," and "the arrangement of musicians with respect to the horn."[7] In these first decades of sound recording, however, there were no widely accepted objective standards to validate and compare reproduction quality. In the absence of fixed parameters or agreed recording standards, different records and competing technologies were judged according to a loose and changing set of aesthetic and technological criteria. These were based on a combination of technological advances; the faith that listeners, engineers, and musicians invested in their devices; and shifting expectations of, and preferences for, certain sonic qualities. In short, although, as Jonathan Sterne notes, the concept of recording "fidelity" already entered the discourse on sound reproduction at this early stage, it initially remained quite ill-defined.[8]

On the one hand, many early accounts of sound recording seem to presuppose an ontological difference between original and copies and cast the recording device itself as a "vanishing mediator."[9] In the 1880s, Sterne describes, Bell's associate Tainter already classified the noises of recording and reproduction devices as "external" to, and separate from, the reproduced sound itself. To achieve "a kind of acoustic transparency in sound reproduction," the separation between this internal sound and external noise ideally meant that "the medium would disappear, and original and copy would be identical for

[5] Gelatt, *Fabulous Phonograph*, 34–35.
[6] Marsha Siefert, "Aesthetics, Technology and the Capitalization of Culture: How the Talking Machine Became a Musical Instrument," *Science in Context* 8, no. 2 (1995): 421.
[7] Thompson, *Soundscape*, 294.
[8] Sterne, *Audible Past*, 218.
[9] Sterne, *Audible Past*, 218.

listeners."[10] Given the many noises that inevitably accompanied early sound reproductions, however, this ideal transparency could only be glimpsed by training listeners to "separate foreground and background sounds" and thereby place "the device as somehow outside the universe of sound reproduction."[11] Fueled by this belief in absolute mediatic transparency, the idealized separation between internal sound from external noise became one of the key narratives in audio engineering, communication technology, and information theory.[12]

On the other hand, if any concept of fidelity were to hold in these first decades of phonographic reproduction, a more flexible attitude toward accuracy was required. Subjective aesthetic considerations would have to be privileged over technological standards and the ideal of transparency. Indeed, in the first few years, when the phonograph was still primarily used to record speech for business purposes, fidelity simply referred to the intelligibility of recorded words.[13] Only from the 1890s onward, as the technology was increasingly marketed as a means of recording and playing music, did fidelity came to encompass aesthetic preferences too. These preferences were not classified according to any standardized system of reference but established through more or less subjective comparisons among different recordings and technologies. Although the concept of fidelity fundamentally presupposed the possibility of similitude between originals and copies, achieved by eliminating every acoustic trace of the medium itself, attempts to establish fidelity in practice did not rely on such objective comparisons of recordings with external "reality" (the "original" sound, prior to recording). They depended instead on subjective comparisons among the sound qualities of different recordings and technologies, based on loose sets of technological, social, and aesthetic values.

This changed with the introduction of electricity. Whereas acoustical recording turns sound waves directly into mechanical vibrations, electrical recording first transduces sounds into an electrical current before they are etched onto the recording surface, and again before they are sent to the loudspeaker. Although its principles were first conceived and patented in 1903, more than a decade had passed before reliable microphones and amplifiers developed in the context of experiments with telegraph and radio technology

[10] Sterne, *Audible Past*, 256.
[11] Sterne, *Audible Past*, 256, 258. Emphasis in original.
[12] In marketing terms, this idea of mediatic transparency took the form of the "tone test," which "equat[ed] phonographic recordings with live performances of music" in front of a live audience to convince listeners that the two were impossible to tell apart. Emily Thompson, "Machines, Music, and the Quest for Fidelity: Marketing the Edison Phonograph in America, 1877–1925," *Musical Quarterly*, 79, no. 1 (1995): 132.
[13] Thompson, "Machines," 137.

during World War I. Although the fundamentals of the procedure remained essentially the same, the electrical recording systems that engineers at Bell Labs began developing from 1919 onward "eliminated," as Susan Schmidt Horning emphasizes, "most [ . . . ] of the major problems associated with acoustical recording and reproduction."[14] What is more, the novel use of microphones meant that musicians no longer had to cluster closely around the recording horn, amplifiers allowed recorded sounds to be reproduced much more loudly, and new electrical recording heads allowed for the reproduction of much greater timbral detail.

Still, despite the clear advantages of electrical recording, objections were also raised to it. Adversaries, Edison himself among them, argued that the transduction of sound waves into an electrical current puts an unacceptable distance between the sound source and recording device. As Schmidt Horning puts it, each electrical component in the reproduction chain "introduce[s] an inherent coloration, or distortion" of the output signal.[15] Because microphones, cables, amplifiers, plugs, and loudspeakers change the reproduced sound, the opponents of electrical recording held, they were all detrimental to the purity of the reproduction. Whereas analog recordings often lacked something (low and high frequencies, for instance) and suffered from ulterior interferences such as needle scratch and surface noise, electrical recordings gained something. Unlike the "external" noises of analog machines, the transduction from sound waves to electricity (and back) makes the recording more susceptible to distortion and noise occurring in the transmission channels between input and output, which directly changes the sound itself.

The technological origin of electrical signal processing in telegraphy, telephony, and radio proved significant for attempts to deal with these side effects. After all, in the years prior to the development of electrical recording equipment, these new and improved communication technologies had already been confronted with "a whole new category of noises that originated in electric systems."[16] As electrical set-ups slowly became standard in recording studios, some of the methods, procedures, and standards that had already been developed in the context of telegraph, telephone, and radio began to be applied to sound recording. Because of this, the introduction of amplifiers, microphones, transistors, and cables (and the electrical noise they produce), was accompanied by the application and further development of new

[14] Susan Schmidt Horning, *Chasing Sound: Technology, Culture and the Art of Studio Recording from Edison to the LP* (Baltimore: John Hopkins University Press, 2013), 35.
[15] Schmidt-Hornig, *Chasing Sound*, 99.
[16] Wittje, *Age of*, 203.

methods of containing and reducing the side effects of these noisy transmission channels.

As Mara Mills has shown, it was with the development of telephone and radio networks in the 1910s and 1920s that "electrical 'perturbations' in the atmosphere and in vacuum tubes became known as 'noise'" and the "concept of noise as 'unwanted' sound [ . . . ] entered the scientific lexicon."[17] As a result, the concept of noise itself was transferred—or rather transduced—from the acoustic to the electrical domain. Just as Tainter classified the noise of the apparatus itself as external to recorded sound, telephone engineers began to regard static electrical noise on telephone lines as "intrinsic to the medium but extraneous to the signal."[18] This redefinition made it possible to develop what earlier attitudes toward sound quality and recording accuracy had lacked: measurable scientific standards enabling objective comparison. By expressing the amplitude of the transmitted signal in relation to the amplitude of the background static, radio engineers formalized noise's influence in terms of the "signal-to-static ratio."[19] Having codified noise as a measurable relation, they subsequently began developing systems to reduce it. As these attempts progressed though the 1920s, the very definition of noise itself gradually widened to include an increasing number of interferences, distortions, and noises that occur along the many transmission channels of communications systems.[20] By the time that the signal-to-static ratio was rebranded a "signal-to-noise ratio" in the 1930s, it had become the univocally accepted standard for expressing the background noise level of any system.

Applied to sound reproduction technology, these objective standards enabled more efficient types of noise reduction. Most of these were inserted into the reproduction chain to reduce noise levels of already recorded material upon playback. In the long run, however, the development of filters that deal specifically with noises in targeted frequency ranges or "bands" was more important. In the 1930s, it was discovered that high frequency noise can be reduced by increasing a sound's amplitude (or volume) upon recording. In this way, the signal could be made louder than the surface noise. On playback, when the original amplitude values are restored, the signal covers or

[17] Mara Mills, "Deafening: Noise and the Engineering of Communication in the Telephone System," *Grey Room* 43 (2011): 123.
[18] Mills, "Deafening," 123.
[19] Schwartz, *Making Noise*, 17–18.
[20] Wittje, "Concepts," 19. Wittje cites engineer Harvey Fletcher, writing in 1929: "When transmitting speech or music either directly to an audience in a large hall or over an electrical system, such as a radio or a telephone system, there is always an interference to the proper reception of such speech and music, due to other sounds being present. These extraneous sounds which serve only to interfere with the proper reception are designated by engineers as 'noise.' With such a designation, the sound may be either periodic or non-periodic as long as it is something that would be better eliminated." Fletcher in Wittje, *Age*, 203–204.

"masks" the noise, which becomes practically inaudible. This method of noise reduction is called "pre-emphasis/de-emphasis." As we will see, when high-frequency noise became increasingly problematic in the era of magnetic tape recording after World War II, it would go on to inspire the most effective noise reduction systems prior to digital recording.

With the transition to electrical recording, and introduction of quantitative measures for standardizing sound recording, transmission and reproduction, the concept of sound fidelity lost some of its most subjective tendencies. The standardization of properties including amplitude levels, frequency response, and dynamic range, combined with newly verifiable and universally accepted standards such as the signal-to-noise ratio, allowed for the more or less exact quantification of sound quality and noise levels. Nevertheless, the idea that sound media are "vanishing mediators" and more relativist interpretations of fidelity persisted alongside this quantitative turn. The newly developed standards supposedly provided more objective measurements of a recording's faithfulness to the original sound source, which was conceived of as existing prior to recording and outside of the sphere of representation. And yet the flexibility introduced by microphones, amplifiers, and devices such as equalizers also opened up a whole new domain of sonic manipulation that "changed the concept of 'original' or 'authentic' performance" altogether.[21] On this basis, Michel Chion argues that these new technologies inspired a more objective concept of "definition," not "fidelity." Whereas the subjective ideal of "fidelity" retained ideals of absolute transparency, the concept of "definition" assesses the quality of a sound reproduction in purely technical terms, regardless of either subjective aesthetic judgements or idealistic callings.[22]

Recording definition essentially describes the extent to which a signal can be differentiated or picked up from amid background noise. Although, as Chion points out, "high definition" is often "(mistakenly) taken as proof of high fidelity," a system's definition is based not on aesthetic considerations or a comparison with the "original" sound source, but on the objective and verifiable parameters of frequency response and dynamic range.[23] Frequency response describes the range of frequencies, expressed in hertz, that a system can reproduce without distortion. Whereas Edison's acoustical Diamond Disc could reproduce frequencies from about 1,000 to 3,000 Hz, the frequency response of a Compact Disc ranges from 20 Hz to 20 kHz (20,000 Hz).[24]

[21] Schmidt-Horning, *Chasing Sound*, 6.

[22] Michel Chion, *Audio-Vision: Sound on Screen*, trans. Claudia Gorbman (New York: Columbia University Press, 1994), 98.

[23] Chion, *Audio-Vision*, 98.

[24] *Encyclopedia of Recorded Sound*, Second Edition, ed. Frank Hoffmann (New York: Routledge, 2005), s.vv. "Tone Tests," "Compact Disc," 2005.

As such, it covers the full, standardized range of human hearing. Dynamic range denotes the difference in amplitude (volume) between the softest (weakest) and loudest (strongest) signal, expressed in decibels. Weak signals do not carry much energy and are easily drowned out by background noise. Accordingly, dynamic range is directly related to noise levels, and maximum dynamic range equals maximum signal-to-noise ratio. Whereas the range of Berliner's earliest gramophone records "did not exceed 6 dB," a regular CD has a signal-to-noise ratio of 96 dB.[25]

The concept of definition does not have the idealistic undertones that attach to fidelity. It presupposes neither the possibility of a "vanishing mediator" nor an intrinsic relation between the output signal and some "original" sound, existing prior to recording. Instead, "definition" relies on technologically verifiable and objectively comparable standards. Whether a wider frequency response or larger dynamic range are considered aesthetically desirable depends on relative and contextual factors, such as the occasion and location of playback, a listener's attention, and the subjective preferences of musicians, engineers, and audiences. Given that such factors not only differ per person, group, and context but also shift over time, efforts to further improve sound technology from the age of electrical recording onward did not seek to cater to flexible aesthetic preferences. Instead, the primary aim behind attempts to reduce noise levels, extend dynamic range, and enlarge the frequency response was to increase overall (objective, measurable) sound definition.

## Magnetic Tape and Noise Reduction

By the 1930s, when electrical recording techniques were well established, FM radio achieved levels of noise suppression that left most disc recording far behind.[26] Around the same time, the movie industry began using pre-emphasis/de-emphasis noise reduction on movie soundtracks. Full frequency range recording was subsequently developed for military purposes in World War II. This allowed for the last major achievement of the era of disc recording: the microgroove disc. More commonly known as the LP, the microgroove disc

[25] *Encyclopedia of Recorded Sound*, s.v. "Signal-to-Noise Ratio," As I explain further on in this chapter, the objective dynamic range and signal-to-noise ratio of a digital recording can differ from the subjectively perceived dynamic range. With the use of "dithering" and "noise shaping," the perceived dynamic range of a 16-bit digital recording can be "about as great as 115 dB." Bob Katz, *Mastering Audio: The Art and the Science* (Oxford: Focal Press, 2002), 51.

[26] Mischa Schwartz, "Improving the Noise Performance of Communication Systems: Radio and Telephony Developments of the 1920s," *IEEE Communications Magazine* 47, no. 12 (2009): 20; Millard, *America*, 277.

was developed at Columbia Records from 1945 onward and realized in 1948. With the "amplifier, record material, shape of the groove, cartridge and stylus, method of recording, [and] turntable drive" all radically updated, the LP constituted a complete overhaul of every aspect of the reproduction chain.[27] The influence of noise and distortion at each link in the chain was drastically reduced, achieving the highest sound definition to date. Vinyl LPs only reached their full potential, however, following another major innovation: magnetic tape recording.

Alongside electrical recording, the introduction of magnetic tape represents the most important revolution in recording practices before the digital age—not least because it enabled the most effective analog technological noise reduction systems ever developed. Although Danish inventor Vladimir Poulsen had already developed the basic principles of magnetic recording with his "telegraphone" of 1899, its recordings were very quiet.[28] Given that microphones and amplifiers had not yet been developed, his invention remained unsuccessful. Only when technological amplification became available in the 1920s did experimentation with magnetic tape begin in earnest. The German "magnetophone" was introduced in 1935 and improved for propagandistic purposes during World War II. Its most significant improvement was the addition of an "ac-bias"-signal: "a high-frequency alternating current," playing alongside the recorded audio, that helped to correct major nonlinearities in the magnetic medium.[29] This significantly decreased tape noise and significantly increased definition. When the German magnetophones were discovered by Allied forces toward the end of the war, they were brought back to the United Kingdom and United States for further improvement.[30] Only a decade later, in the mid-1950s, the transition from disc to tape was all but complete.[31]

In 1952, when the technology was still very new, Read described the operations of "magnetic recording machines" in the following way:

[27] Quote by engineer Peter Goldmark, who developed the LP, in Mark Coleman, *Playback: From the Victrola to MP3; 100 Years of Music, Machines, and Money* (Cambridge, MA: Da Capo Press, 2003), 28.

[28] N. Katherine Hayles notes that, even prior to Poulsen, "as early as 1888 Oberlin Smith, at one time president of the American Society of Mechanical Engineering, proposed that sound could be recorded by magnetizing iron particles that adhered to a carrier." N. Katherine Hayles, "Voices Out of Bodies, Bodies Out of Voices: Audiotape and the Production of Subjectivity," in *Sound States: Innovative Poetics and Acoustical Technologies*, ed. Adalaide Morris (London: University of North Carolina Press, 1997), 76.

[29] *Encyclopedia of Recorded Sound*, s.v. "Bias."

[30] Read, *Recording*, 190.

[31] By the mid-1950s, writes Beverly R. Gooch, "magnetic audio recording had completely revolutionized the record and broadcasting industry. All records were mastered on tape, and radio broadcasters were exclusively using tape as a time-delay and programming tool." Beverly R. Gooch, "Building on the Magnetophon," in *Magnetic Recording: The First 100 Years*, eds. Eric D. Daniel, C. Denis Mee, and Mark H. Clark (New York: The Institute of Electrical and Electronics Engineers, Inc. 1999), 90, 72–91.

> A magnetic material, such as wire or tape, is drawn past a recording head. As it passes through the head, the material becomes and remains magnetized. The amount of magnetization remaining in the material at each instant is governed by the impressed signal upon the recording head. In playing back, the magnetized material is drawn past a playback head. The varying magnetization which remains in the material induces corresponding voltages in the coil of the playback head.[32]

The advantages of this procedure were manifold: first, because tape recording avoids direct contact between mechanics (the recording head) and surface (magnetized tape), the persistent problem of surface noise all but disappeared; second, the maximum frequency response was no longer "limited by the inertia of mechanical parts," as Schmidt Horning puts it, and the maximum dynamic range no longer "limited by the dimensions of the groove."[33] Magnetic recording meant another big leap in flexible recording practices. With an uninterrupted thirty-minute run, sessions could last longer. New techniques became available, including splicing and editing the tape; changing recording or playback speed; reversing sounds; and creating otherworldly echo effects. Most strikingly, magnetic tape provided efficient ways of multitrack recording or "overdubbing," which enables artists to build up complex musical pieces out of recordings made at different times. Various musical parts and takes, taped separately and at different times, can be combined, layer by layer. Overdubbing unlocked a whole new suite of creative studio practices. The musical possibilities to which it has given rise have defined avant-garde and popular music alike from the 1950s and 1960s onward, from electronic compositions such as Karlheinz Stockhausen's "Gesang der Jünglinge" ("Song of the Youths," 1955–1956) to Brian Wilson's four-minute "pocket symphony" that is the Beach Boys' "Good Vibrations," which was constructed out of more than ninety hours of recorded tape.[34]

As always, however, these new technologies and new possibilities also introduced new noises, interferences, and distortions, and magnetic tape recording's greatest achievement—significantly increased frequency response—risked becoming a problem. With this dramatically improved frequency range, a new type of high-frequency noise called "tape hiss," the result

[32] Oliver Read, *The Recording and Reproduction of Sound: A Complete Reference Manual on Audio for the Professional and the Amateur*, Second Edition (Indianapolis: Howard W. Sams, 1952), 181.

[33] Schmidt Horning, *Chasing Sound*, 106.

[34] Karlheinz Stockhausen, *Gesang der Jünglinge*, track 4 on *Elektronische Musik 1952–1960*, Stockhausen-Verlag, 2001, compact disc; "Good Vibrations," track 6 on The Beach Boys, *Smiley Smile*, Capital Records/Brother Records, 1994, compact disc; Tom Pinnock, "The Making of. . . . the Beach Boys' 'Good Vibrations,'" *Uncut*, June 8, 2012, accessed September 2, 2019, https://www.uncut.co.uk/features/the-making-of-the-beach-boys-good-vibrations-34867.

of magnetized particles on the tape, became increasingly audible as well. This was a real nuisance, especially during low-volume passages in which the signal does not cover or "mask" the hiss. Sound engineers had already confronted the surface or scratch noise of wax cylinders, shellac, and vinyl records, and the transmission noises caused by electrical recording. Now tape hiss and other noises that became audible with increased definition required a renewed attempt to save sound reproduction from its own internal enemies. This challenge was taken up by engineer Ray Dolby. Combining advances in electrical signal processing of the 1930s with magnetic tape's increased recording flexibility, he developed the most sophisticated and effective noise-reduction technology of the analog era.[35] His system implemented a "dual-ended" noise-reduction process, based on the earlier pre-emphasis/de-emphasis principle.

"Single-ended" systems work their magic at either the beginning or end of the recording chain, reducing noise before the input signal is recorded or before the output signal is played back. To confront noises that appear prior to recording (such as those produced by microphones, amplifiers, effect modules, cables, or electronic musical instruments), one can install an adaptive filter. By analyzing and processing the dynamic range and frequency spectrum of the incoming signal, such filters reduce noise in specific frequency bands whenever the signal becomes weaker than the background noise. Similarly, so-called noise gates, which cut off frequencies below or above a given threshold, were developed in the late 1960s.[36] At the other side of the chain, filters can be applied that partly remove the noise of playback media or even reduce noise on already recorded material. Like adaptive filters at the front of the chain, these devices filter noise by attenuating specific frequency bands. More recent digital technologies can even analyze and subtract the frequency spectra of certain noises from a sound file and clean up recordings that have been heavily affected by noise.[37]

Contrary to these single-ended approaches, "dual-ended" systems intervene before both recording and playback. The basic principle of pre-emphasis/de-emphasis is this: first, it boosts ("emphasizes") the signal's amplitude, providing it with enough energy to "mask" the noise. At playback, or during the mastering process, the original amplitude levels are restored

[35] Mark H. Clark, "Product Diversification," in *Magnetic Recording: The First 100 Years*, eds. Eric D. Daniel, C. Denis Mee, and Mark H. Clark (New York: The Institute of Electrical and Electronics Engineers, 1999), 94.

[36] David Miles Huber and Robert E. Runstein, *Modern Recording Techniques*, Seventh Edition (Oxford: Focal Press, 2010), 517.

[37] Huber and Runstein, *Recording Techniques*, 519.

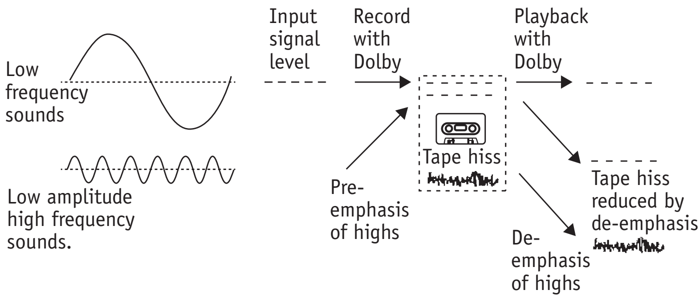


**Figure 1.1** Noise Reduction with Pre-emphasis/De-emphasis. (Reproduced by permission from Rod Nave, "Dolby Noise Reduction," *Hyperphysics*, accessed April 15, 2013, http://hyperphysics.phyastr.gsu.edu/hbase/audio/tape5.html#c1).

("de-emphasized"), thereby lowering the volume of both the recorded signal and background noise: the signal now masks the noise. Dolby refined this process into a more sophisticated procedure called "companding" or "compansion"—compound contractions of "compressing" and "expanding" and "compression" and "expansion" (Figure 1.1). Whereas earlier forms of pre-emphasis/de-emphasis were applied to the entirety of a recording, Dolby operated according to a "principle of least treatment," that is, of avoiding the unnecessary processing of unproblematic sections.[38] Dolby's system therefore uses multiple frequency bands and a dynamic range limiter to ensure that only the passages most affected by noise are processed. Prior to recording, the signal is "encoded" by compressing the dynamic range of specific frequency bands (in the first models, these were primarily the higher regions of the frequency spectrum, where tape hiss is most prominent). By increasing the amplitude of these segments, encoding masks the hiss. "Decoding" happens during playback or mastering. By expanding the dynamic range and restoring original amplitude levels, it reduces noise levels along with the signal, with the result that the noise becomes inaudible.[39]

Between the mid-1960s and 1990, Dolby developed multiple generations of his dual-ended noise reduction system, both for professional studios (Dolby A and SR) and simplified versions for consumer cassette tapes (Dolby B, C, and S). The first system consisted of four frequency bands and reduced noise, especially in the higher frequency ranges, with a maximum of 10 dB. Dolby Laboratories claims that their final analog system, the Dolby SR of 1986,

[38] "Dolby B, C, and S Noise Reduction Systems: Making Cassettes Sound Better," *Dolby Laboratories*, 2001, accessed January 3, 2013, www.dolby.com/uploadedFiles/English(US)/Professional/Technical_Library/Technologies/Dolby_A-type_NR/212_Dolby_B,_C_and_S_Noise_Reduction_Systems.pdf.

[39] Huber and Runstein, *Recording Techniques*, 515–516.

reduces noises and "other low-level disturbances" over the spectrum "by as much as 25 dB."[40] This effectively brings all noise below the level of audibility to produce what they call a "remarkable clarity of reproduction."[41] According to Dolby's brochure, the goal of this noise reduction is to approximate "an ideal audio device or system [that] would impose no audible limitation on the signal passing through it."[42] Evidently, almost exactly one hundred years after the development of the graphophone, Tainter's ideal of the vanishing mediator was still very much alive. And with the transition to digital technology from the 1970s onward, this ideal of a fundamentally inaudible medium producing entirely noise-free reproductions and absolutely clear signals seemed to come even closer. Even the digital revolution, however, did not escape the continuous occurrence and reoccurrence of noise in sound reproduction.

## Separating Signal and Noise: Digital Recording

By the time Dolby Laboratories launched its final analog noise-reduction system in the spring of 1986, the market for sound-reproduction technologies looked very different from when Ray Dolby had demonstrated his first system in 1965. Four years earlier, Philips and Sony presented the first commercial digital sound carrier, the Compact Disc. The culmination of intensive research going back to the mid-1920s, it was a final step in the slow but steady takeover of digital sound technologies that began in the 1970s, gained momentum in the 1980s, and was more or less completed by the 1990s.[43] The first and still most commonly used principle in converting analog sound into digital signals, pulse code modulation (PCM), was patented in 1926 and further developed in 1937 in the context of telephone engineering. Based on "the concept that a continuous signal could be reconstructed from isolated samples and that these samples could be approximated by discrete numbers," PCM turns sound "into a pulsating electric current that is measured and expressed as a binary code of digits."[44] This binary code, representing the current's amplitude values, is inscribed on a hardware medium: as pits in a surface (as with

[40] "Dolby® SR. Dolby® Spectral Recording," Dolby Laboratories, 1987. Accessed January 14, 2013, www.dolby.com/uploadedFiles/English_(US)/Professional/Technical_ Library/Technologies/Dolby_Spectral_Recording_(SR)/33_SpectralRecordingPaper.pdf, 3-5.
[41] "Dolby® SR," 3–5.
[42] "Dolby® SR," 2.
[43] David Morton, *Off the Record: The Technology and Culture of Sound Recording in America* (New Brunswick: Rutgers University Press, 2000), 172. For a concise history of the development and introduction of the Compact Disc, see chapter 6 of Greg Milner, *Perfecting Sound Forever: An Aural History of Recorded Music* (New York: Faber and Faber, 2009).
[44] Millard, *America*, 348.

the CD), spots of magnetic flux (as with digital audio tape, or DAT), or any other binary codification system. At playback, the code is read by a laser or magnetic tape head, translated back into electrical pulses, and transduced back into sound waves.

During World War II, these principles were put to practical use by Bell Labs as a means of codifying telephone messages between the United Kingdom and United States. Extensive research into the possibilities of digital signal processing began after Claude Shannon (also at Bell Labs) formalized and refined the theoretical principles of information theory in 1948. It continued throughout the 1950s and 1960s. The first experiments in digitally synthesizing complex sounds from scratch were conducted in 1957.[45] Just a few years later, in 1962, Bell Labs installed the first digital signal transmission system. The age of digital *recording* subsequently started in 1967 at the NHK Technical Research Laboratory in Japan.[46] New and better digital recorders were rapidly developed through the 1970s, and the first commercial digital synthesizer, the Synclavier, appeared halfway through the decade. "By the beginning of the 1980s," Thomas Fine writes in his brief history of digital sound technology, "all major record companies had embraced digital recording in one form or another."[47] This cleared the way for a fully digital sound carrier: the CD. Although tape was widely used well into the 1990s, in principle, the entire chain of recording, transmission, and reproduction could now be carried out digitally. The age of digital sound went into full swing.

In many ways, the advantages of digital sound recording echo those of electrical recording and magnetic tape: increased definition and flexibility in recording. From the 1990s especially, the development of user-friendly digital recording hardware and software unlocked the potential of digital sound processing for an increasingly large group of people, professionals and amateurs alike. Furthermore, when the correct conditions are observed, "digital technology can create any number of generations of perfect (noise-free) clones of an original recording."[48] This means that, in principle, every copy of a digital recording is exactly the same as the master recording. The most important promise of digital technology, though, was its potential for highly increased sound definition. Indeed, digital sound technology allows for a frequency response and dynamic range that outreach even the most advanced analog

[45] Curtis Roads, *The Computer Music Tutorial* (Cambridge, MA: MIT Press, 1996), 87.
[46] Thomas Fine, "The Dawn of Commercial Digital Recording," *ARSC Journal* 39, no. 1 (2008): 3. According to Fine, the first commercially available digital recording made with this Japanese prototype was a cover of the Beatles' hit "Something" by jazz musician Steve Marcus, recorded in 1970 and released in January 1971. Fine, "Dawn," 4.
[47] Fine, "Dawn," 14.
[48] Roads, *Computer Music*, 21.

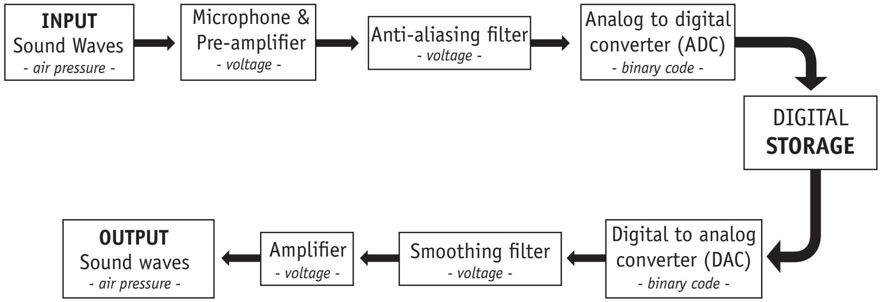


**Figure 1.2** Digital Recording and Reproduction Chain.

system. A CD can easily reproduce a frequency range from well below 20 Hz to over 20 kHz, and a dynamic range of at least 96 dB.

Digitizing sound, however, is a complex procedure that requires specialized engineering skills and the correct use, adjustment, and calibration of highly sensitive equipment. Having been captured and turned into electrical waveforms by a microphone, sound waves are subsequently converted into binary code by an analog-to-digital (A/D) converter (Figure 1.2). The A/D-converter cuts, or "samples," the signal into many thousands of discrete pieces of time. For every one of these samples, the voltage level of the electrical impulse (which corresponds to the amplitude level of the original sound wave) is measured and "quantized."[49] This means that the measured voltage levels are translated into binary code. This twofold operation of sampling bits of time and quantizing voltage or amplitude values constitutes the basis of sound digitization. Each sample contains a binary number representing the signal's amplitude value at the moment of measurement. Combined in sequence, these samples (in the case of a normal CD, 44,100 samples per second) represent the original waveform.

At playback, this process is reversed: the binary numbers are read by a digital-to-analog converter (DAC) that translates binary numbers back into voltage levels, or digital code back into an electrical current, which is subsequently transduced into sound waves. Contrary to popular belief, the parts of the signal that fall "in between" the samples are not lost or unrecorded. Under the right circumstances, the DAC (aided by a "smoothing filter") connects the dots between the discrete samples, restoring those missing parts of the signal that fell "between the samples."[50] Provided that the ideal conditions are met,

[49] Francis Rumsey and Tim McCormick, *Sound and Recording*, Sixth Edition (Oxford: Focal Press, 2009), 203.
[50] Roads, *Computer Music*, 27.

then, the digital procedure is able to reproduce the input signal exactly. The physical realization of this theoretical perfection is limited, however, by the varying extent to which those ideal conditions can be practically established.

With digital media, the fight against the noise of sound media entered a new era. Already in 1951, physicist and cybernetician John von Neumann wrote that "the real importance of the digital procedure lies in its ability to reduce the computational noise level to an extent which is completely unobtainable by any other (analogy) procedure."[51] Although digital computers operate on the basis of analog circuitry, which is just as susceptible to noise and physical interference as any other apparatus, the decisive gesture of digital technology is the symbolical separation of these analog processes from calculations on the computational level. This separation makes it possible to distinguish digital signals from the background noise of physical media much more efficiently than before. Regarding the presence of noise, therefore, von Neumann argues that the fundamental difference between analog and digital technology is not qualitative, but quantitative. Although digital operations are still based on analog processes, noise and signals can be symbolically separated to an extent that is structurally unobtainable using an analog machine.[52] The digital procedure no longer physically inscribes signals on a recording surface but relies instead on measuring signal values and representing them symbolically in binary code. Accordingly, issues such as surface noise, needle scratch, irregularities of the material, tape hiss, and nonlinearity cease being problems. In short, that which had always resisted complete reduction in the analog realm (the physical noise introduced by storage, reproduction, and transmission devices) is almost entirely absent in digital media.

As Bernhard Siegert has argued, just like alphabetic writing, which also uses series of discrete signs to represent a continuous stream of information, digital systems are based on "the filtering out of signals from noise."[53] Whereas the material basis of analog media always affects the inscription, transmission, or reproduction of signals, these discrete writing systems—whether alphabetic or binary—symbolically separate their substrate or base material (paper, pen, ink, plastic, tape, silicon circuitry) from written signs (letters or

[51] John Von Neumann, "The General and Logical Theory of Automata," in *Collected Works*, Volume 5, ed. A. H. Taub (Oxford: Pergamon Press, 1963), 295. Von Neumann specifically discusses the difference between analog *computers*, which are based on processing "the intensity of an electrical current, or the size of an electrical potential, or the number of degrees of arc by which a disk has been rotated," and digital *computers* that "represent [ . . . ] numbers as aggregates of digits." Regarding the issue of noise, this structural difference between analog and digital *computers* can be extrapolated to the difference between analog and digital machines in general. Von Neumann, "Theory," 293–294.

[52] Von Neumann, "Theory," 295.

[53] Bernhard Siegert, *Cultural Techniques: Grids, Filters, Doors, and Other Articulations of the Real*, trans. Geoffrey Winthrop-Young (New York: Fordham University Press, 2015), 30.

numbers). In this way, the noise of material transmission channels does not interfere with the message. The order of digital signals returns to this discrete logic of the written sign: in digital media, noise reduction is not an additional filter applied before, during, or after the reproduction process. Instead, the parting of signal from noise—which takes the form of a symbolic separation of analog circuits from digital computation—is constitutive of the digital as such. Hence, Siegert concludes, in "the order of digital signals," the logic of separating signals from noise "becomes nothing less than systemic."[54]

Given its systemic separation of signal and noise, digital technology is often considered the final word on noise reduction. "Here at last," Andre Millard writes with a pathos that echoes the jubilant enthusiasm of early advocates of digital technologies,

> was a system of recording in which there was no extraneous noise: no surface noise of scratches and pops, no tape hiss, and no background hum. The compact disc has a signal-to-noise ratio of 96 dB, which in effect makes it noiseless recording.[55]

The physical reality of digital signal processing, however, is not as flawless as the idealized theory might suggest.[56] The mathematical models at the basis of digitization might indicate the possibility of perfect reproduction, but its implementation in physical hardware poses specific problems. For the present discussion, the most important of these are "aliasing" and "quantization errors." Both issues derive directly from the basic principles of digitization.[57]

The issue of aliasing touches upon one of the cornerstones of the digital procedure: the sampling theorem (also called the "Nyquist" or "Shannon-Nyquist" theorem, after Harry Nyquist, who first described it in 1928, and Claude Shannon, who formalized the theorem in 1949). The sampling theorem defines the minimum number of samples necessary

[54] Siegert, *Techniques*, 30.

[55] Millard, *America*, 353. Compare the enthusiasm of engineer Ken C. Pohlmann, who wrote the following at the dawn of the age of digital sound recording in 1985: "Now the wait is over. With digital music one can at last listen to playback and begin to feel as if one is there—at the performance. High fidelity will have to be redefined as higher fidelity." Ken C. Pohlmann, *Principles of Digital Audio*, First Edition (New York: McGraw-Hill, 1985), 266.

[56] Indeed, Roads emphasizes physical implementations of digitization: "In contrast to the myth of 'perfect reconstruction' which pervades the mathematical theory of signal processing, the actual quality of all analysis-resynthesis methods is limited by the resolution of the input signal and the numerical precision of the analysis procedures." Roads, *Computer Music*, 273.

[57] Besides aliasing and quantization errors, probably the most well-known problem causing digital distortion is "jitter," which is the result of temporal irregularities in the sampling process. When the sampling of bits of time is not carried out in strictly regular intervals, random noise and periodic distortions result, causing a considerable loss in sound definition. A well-calibrated, high-quality sample clock minimizes the risk of jitter. For more on jitter and digital clocking, see Owen Marshall, "Jitter: Clocking as Audible Media," *International Journal of Communication* 13 (2019): 1846–1862.

for reproducing a given frequency spectrum without errors occurring. It stipulates that the minimum sample rate (the number of samples per second) must be at least twice as high as the highest frequency (number of oscillations per second) of the reproduced signal. This means that a system must take at least two samples for each full wave or each “cycle” of the highest frequency if it is to adequately represent a signal and turn it back into a faultless sound wave. In accordance with the sampling theorem, the standard sample rate of a CD (44,100 samples per second) can encode a frequency spectrum up to roughly 22 kHz. It is no coincidence that this slightly exceeds the standardized upper threshold of human hearing (20 kHz).

If a signal is digitized with a sample rate of less than twice the highest frequency, then the sampling process will be unable to faithfully represent frequencies above the “Nyquist threshold” of half the sample rate. These frequencies will “fold over” and appear in the reconstructed sound as a lower frequency, thus introducing “alias frequencies” that were not present in the input signal (Figure 1.3). To prevent the appearance of these fold-over frequencies, an anti-aliasing filter cuts off all frequencies above the Nyquist threshold—filtering out everything above 20 kHz, for instance. For reasons that will be discussed in more detail in chapter 3, however, real-time filters cannot abruptly cut off a signal at some arbitrary threshold. They need time to process the signal, causing a delay that, however minuscule, allows part of the frequency spectrum to sift through. In practice, the sample rate should therefore be slightly higher than the Nyquist threshold, permitting the filter to have a gradual cut-off slope, as shown in Figure 1.4.

Improving anti-aliasing filters and other strategies that prevent aliasing has been a major concern for engineers ever since the introduction of digital recording. Despite these efforts, “filter effects are unavoidable to some

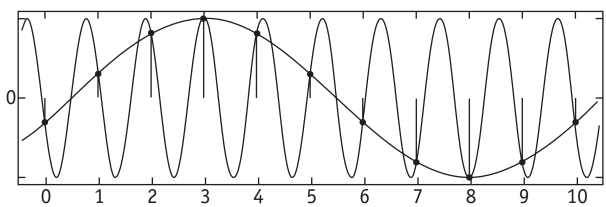


**Figure 1.3** Aliasing. A sinusoidal wave is sampled at a rate below the Nyquist-threshold of twice its frequency. Based on samples 1 to 10, the digitization procedure produces a much lower alias frequency. (Courtesy of Moxfyre, “AliasingSines.svg” *Wikimedia*, accessed November 27, 2019, https://commons.wikimedia.org/wiki/File:AliasingSines.svg).

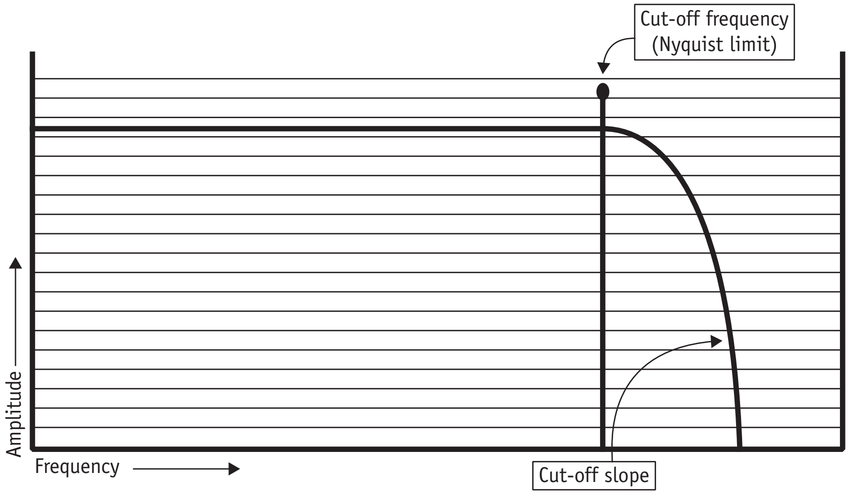


**Figure 1.4** Cut-off Slope. To prevent aliasing, the frequency range of a digitized signal should be cut off at the Nyquist-limit of twice the maximum frequency. Because a hard limit will allow part of the spectrum to sift through, however, the sample rate should be slightly higher, to allow for a gradual cut-off slope.

extent"—even when much higher sample rates are used.[58] Because the sample rate determines the width of the frequency spectrum that a system can reproduce, it is directly related to its maximum frequency response. And because filters with abrupt cut-off frequencies remain physically impossible, measures to prevent aliasing always introduce some effects in the reproduced signal, however slight. The Nyquist theorem, therefore, indicates a first reason why theoretically perfect digitization is always compromised in physical reality. Quantization errors, however, produce an even more fundamental limitation.

Quantization means the measurement of voltage levels in a sampled signal and translation of the results into binary "words," which numerically represent the original levels. Each binary digit (1 or 0) in such a word constitutes one "bit" of information. The number of available bits, called "bit depth" or "sampling precision," designates the possible length of each word. Hence, greater sampling precision equals more available bits and longer possible word lengths. And because longer words allow voltage values to be stored more accurately, the precision of the quantization process increases with each available bit. In the case of sampling, a signal can be fully restored as long as the sampling theorem is not violated. With quantization, however, "the values of the sampled signal," Roads describes, "cannot take on any conceivable

[58] Rumsey and McCormick, *Sound and Recording*, 216.

[voltage] value. This is because digital numbers can only be represented within a certain range and with a certain accuracy."[59] If, as is often the case, the voltage level ideally requires a very long number value to be stored, the converter rounds off the binary word to make the measured value fit the available number of bits. This difference between actual voltage values and approximate quantized values can cause "round-off errors" or "quantization errors."

Like aliasing, quantization errors result in distortion. In the case of high-amplitude signals with a complex frequency spectrum—including many musical signals, like the sound of a full orchestra or a rock band—this error sounds similar to analog noise. As with tape hiss, however, the distortion caused by quantization errors becomes more problematic in the case of very low-amplitude signals, albeit for different reasons. Indeed, with low-amplitude signals, over the course of many consecutive samples, the distortion caused by quantization errors becomes very different from the more or less evenly distributed noise floor of analog technologies. The reason for this is that, when amplitude levels drop, the error no longer translates into random, uncorrelated noise (as besets analog recordings), but into semi-periodic signals that are statistically correlated to both each other and the digitized signal. On top of this, quantization errors can also introduce, as Roads explains, "harmonics [that] may even extend beyond the Nyquist frequency, causing aliasing and introducing new frequency components that were not in the original signal."[60] Usually, these unwanted artifacts of quantization errors are considered a form of so-called harmonic distortion.[61]

Because longer words can store more exact values, the problem of quantization errors and harmonic distortion decreases when the sampling precision is raised. However, more bits require more finite storage space. "In theory," writes Jay Kadis, "the more bits used to encode sample words, the greater our confidence in the accuracy of the measurement and the better the sound quality." In practice, however, "the physical process of conversion limits the real accuracy predicted by theory because real converters fall short of theoretical perfection."[62] Although higher precision rates minimize quantization errors, then, absolute precision is structurally unattainable. Interestingly, however, the solution to the problem of quantization errors and harmonic distortion comes in an unexpected form: noise itself.

[59] Roads, *Computer Music*, 33.
[60] Roads, *Computer Music*, 36
[61] Rumsey and McCormick, *Sound and Recording*, 223.
[62] Jay Kadis, *The Science of Sound Recording* (Oxford: Focal Press, 2012), 145.

## Dithering: Fighting Noise with Noise

At first sight, the practice of "dithering" seems counterintuitive in that it deliberately adds a small amount of noise prior to digitization to counter quantization errors. Dithering works, however, because the distortion caused by quantization errors is statistically correlated. By randomizing this statistical correlation of quantization errors, dither noise cancels out the negative effects of the added noise. With dither, as Figure 1.5 shows, the errors no longer translate into (statistically correlated) harmonic distortion. They are "decorrelated" and become a slight layer of random noise. Instead of petering out into annoying signals, the low-amplitude tones fade smoothly, in Roads's words, "into a bed of low-level random noise."[63] In most cases, especially when the sampling precision is lower than 20 bits (and remember that regular CDs have a bit depth of 16 bits), A/D-converters explicitly add dither noise to signals prior to digitization. In other cases, dithering can "happen naturally by means of the thermal noise within the converters," rendering the addition of noise unnecessary.[64]

Dither noise is not only added at the stage of analog to digital conversion, for sample values need to be requantized during many processes in digital mixing or mastering, for instance while editing, changing volume levels, or adding sound effects. Given the complicated calculations that even the simplest digital processing requires, number values can rapidly expand and introduce new round-off errors. Whereas the CD-standard of 16-bits is still most commonly used for commercially released digital audio, higher precision (up to 24-, 48-, and sometimes even 96-bits) is often used in mixing and mastering to reduce the risks of these requantization errors. These extra bits provide more room (in the form of longer word lengths) for computations without the need for constant dithering. When these high precision master recordings are eventually transferred to the 16-bit standard of CDs or other digital formats, however, they must be requantized (or "truncated") to fit the shorter word length available. This requantization requires dithering.

[63] Roads, *Computer Music*, 37.

[64] Nika Aldrich, "Dither Explained: An Explanation and Proof of the Benefit of Dither," *Cadenza Recording*, April 25, 2002, accessed September 20, 2013, www.users.qwest.net/~volt42/cadenzarecording/DitherExplained.pdf. Mastering engineer Bob Katz nuances Aldrich's statement: "every well-made 16-bit A/D incorporates dither to linearize the signal. If you were lucky enough to have a 20-bit or 24-bit A/D and 24-bit storage to begin with, then dither is probably not necessary during the original analog encoding. Although the inherent thermal noise on their inputs is not shaped to perfectly dither the source, current 20-bit A/Ds self-dither to some degree around the 18–19 bit level because of this basic physical limitation. Similarly, a transfer from typical analog tape probably has enough hiss to self-dither any transfer to 16-bits, as long as there is no digital processing before storage." Katz, *Mastering*, 51.

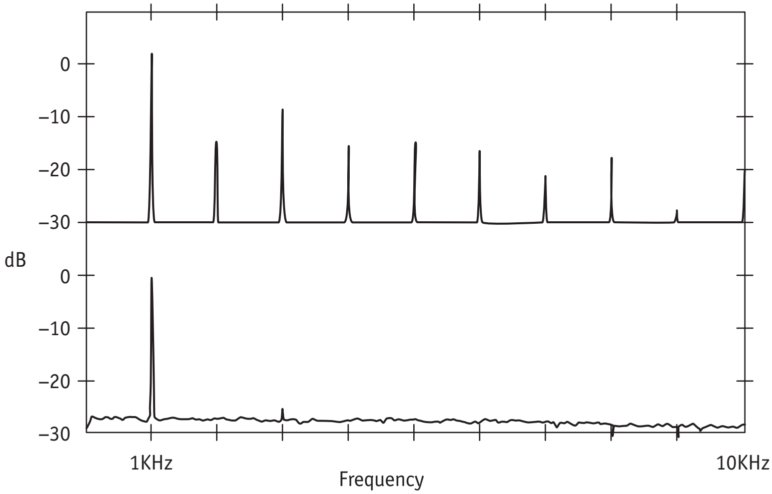


**Figure 1.5** Harmonic Distortion and Dither. The long spike on the left is a single frequency. In the upper diagram, it is quantized without dithering. Harmonic distortion, caused by quantization error, is visible as shorter spikes on the right. Below, the same signal is digitized with dither: the harmonic distortion is gone, and a minor random noise floor can be seen. (Reproduced by permission from Curtis Roads, The Computer Music Tutorial (Cambridge Ma.: MIT Press, 1996), 37).

Just as sample rate determines maximum frequency range, so the number of available bits determines the precision of stored amplitude levels. Bit depth is therefore directly related to maximum dynamic range. The more bits, the more precisely low-amplitude signals are stored and the larger the potential dynamic range. As a rule of thumb, every bit amounts to approximately 6 dB in dynamic range, which is why 16-bit systems (like the CD) boast a maximum dynamic range of 96 dB.[65] In the case of analog media, such as vinyl discs or magnetic tape, the dynamic range is limited by the amount of background noise. This is why effective noise-reduction strategies, such as Dolby's companding system, result in increased dynamic range.

In undithered digital recordings, low-amplitude signals are not affected by background noise (as they would be in the analog scenario), but by quantization errors. This means that the signal-to-noise ratio of analog recordings becomes a signal-to-error ratio in digital recording. When dithering reintroduces noise, this signal-to-error ratio can again be expressed as a signal-to-noise ratio; or, more accurately, as a signal-to-dither ratio. Without

[65] Rumsey and McCormick, *Sound and Recording*, 223.

any dithering, a 16-bit digital recording does not suffer from background noise: the maximum signal-to-noise ratio is 96 dB, which equals the system's maximum dynamic range. Because it does suffer from quantization errors, however, the signal-to-error ratio of the same undithered recording is significantly lower. The addition of dither, in turn, decorrelates the error, but it usually accounts for about 3 to 5 dB of noise. So, the signal-to-dither ratio of a dithered 16-bit recording is somewhere between 91 and 93 dB.

It is important to note that dither does not mask or cover the distortion caused by quantization errors. It does not, in a reversal of the principles of Dolby's noise reduction, use noise to cover unwanted signal. In the case of masking, noise is not removed or eliminated. It is only drowned out and thus rendered inaudible. With dithering, in contrast, the artifacts of quantization error are entirely eliminated. When the original error is decorrelated through the addition of dither, it does not disappear but takes on a radically different shape and ceases to be an error. The error is broken up and redistributed randomly, becoming a different artifact, which closely resembles a minute random noise floor. With the help of additional techniques, this evenly distributed noise floor can subsequently be further reduced or rendered less noticeable.[66] Still, besides reducing the maximum dynamic range by adding this slight noise floor, dither also affects the dynamic range in a more positive way. This is due to a second significant effect, called "stochastic resonance."

Due to the stochastic resonance effect, dither noise pushes very faint signals, which would not otherwise be picked up by A/D-converters, above a minimum threshold of registerability. Because a real digital system can only represent finite value numbers, the values of these very low-amplitude signals normally fall below a minimum threshold of 0.5. Accordingly, systems regularly round them off as 0, which represents "nothing" or, in sonic terms, silence. In other words, this low-amplitude sonic detail is lost entirely. As with quantization errors, more sampling precision (more bits) reduces the pertinence of this problem. In an undithered recording, however, some amount of low-amplitude detail will inevitably be rounded off to zero and thus lost. The addition of a slight layer of dither noise adds some extra energy to these signals and thereby triggers what physicist Luca Gammaitoni calls a "noise activated

[66] The noise floor of dithered recordings can be further reduced using "noise shaping," by means of which, as Rumsey and McCormick write, "noise within the most audible parts of the audio frequency range is reduced at the expense of increased noise at other frequencies." Instead of spreading the noise equally over the signal's entire frequency range, "digital filtering is employed to shape the spectrum of the quantizing noise," thus moving it away from "where the ear is most sensitive and increasing it at the high-frequency end of the spectrum." If necessary, companding techniques or single-ended noise reduction filters can be applied to reduce the noise level even further. Rumsey and McCormick, *Sound and Recording*, 231–232, 236.

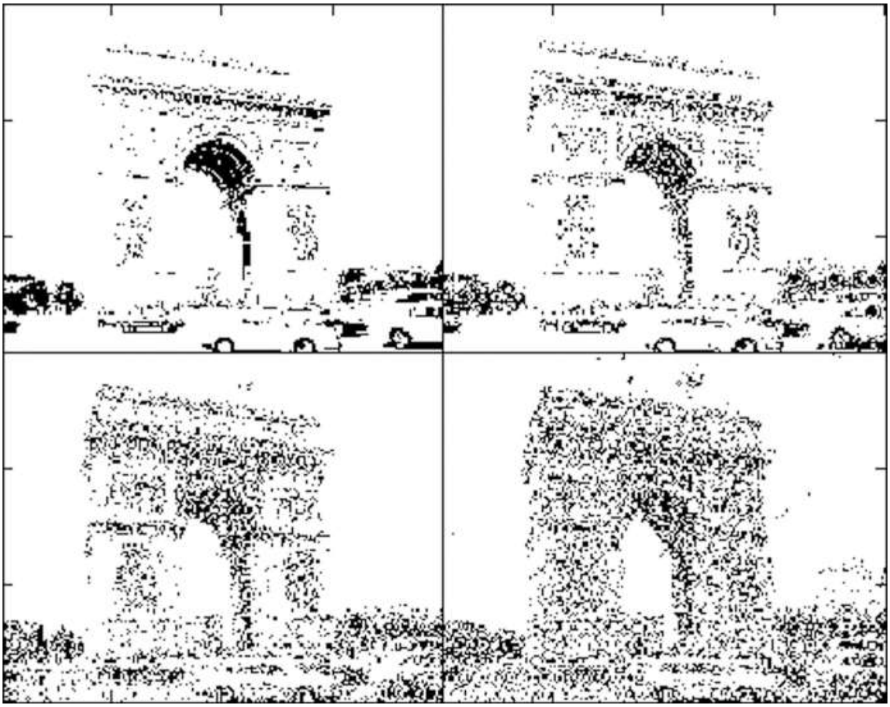

**Figure 1.6** Stochastic Resonance. These four images of the Arc de Triomphe in Paris are rendered with different amounts of visual dither noise. The upper left image is rendered without any extra noise. It has the lowest level of detail. Various levels of noise (in the form of random pixels) are applied to the other three. The image on the lower left seems to have the best resolution, whereas the noise begins to overtake certain details in the image on the lower right. (Courtesy of Jamesvoltage, "Processed image of the Arc de Triomphe," *Wikimedia*, December 5, 2009, https://commons.wikimedia.org/wiki/File:Arcfour2.png, accessed November 12, 2019).

process."[67] It pushes the low-amplitude signals over the minimum threshold of 0.5, thus making sure that their amplitude values are rounded off to 1, not 0, and that the signals are registered by the A/D-converter. The stochastic resonance effect, then, entails the storage of more detailed information, which effectively increases the perceived dynamic range well beyond the maximum signal-to-noise ratio of any undithered recording. Still, as visually illustrated by Figure 1.6, this only occurs within certain limits: a minimal amount of

[67] Luca Gammaitoni, "Stochastic Resonance and the Dithering Effect in Threshold Physical Systems," *Physical Review E* 52, no. 5 (1995): 4691.

added noise pushes faint signals over the threshold of registerability, but too much noise will drown them out again.[68]

The addition of the tiniest amount of noise—some grains of sand in the machinery, as it were—transforms digital systems, allowing them to process signals that would otherwise go unnoticed. Too much noise, however, blocks transmission. It closes the channel. So "how much noise is necessary?"[69] A first step toward answering this question (as posed here by Serres) is hinted at by the history of noise reduction. The continuous reappearance of noise and distortion in ever different forms and places over the course of this history suggests that we should not, as the myth of perfect fidelity would have it, take noise as a byproduct of the recording and reproduction chain—that is, as something to be eliminated, masked, reduced, or filtered out at any cost. Indeed, any seemingly progressive account of sound recording and noise reduction is put in a very different light by the observation that digital technology—despite reducing the material noises of sound reproduction to unprecedented levels—marks the return of noise in the form of deliberately added dither.

Obviously, the search for high definition sound did not cease in the 1980s. It continues to this day. Since the introduction of the CD, new digital carriers and formats like DAT, the Super Audio CD (SACD), and the Minidisc (MD) each promised ever-more splendid sound definition. More recently, the pushback against low resolution file formats such as the MP3, which trade the advantage of smaller file sizes for a loss of definition, has inspired the introduction

[68] According to Pohlmann, the term "dithering" goes back to World War II, when it turned out that analog computers on bomber planes performed "more accurately when flying on board the aircraft, and less well on ground," because, "the vibration from the aircraft reduced the error from sticky moving parts. Instead of moving in short jerks, they moved more continuously." To be able to calibrate the computers while the airplanes were on the ground, "small vibrating motors were built into the computers, and their vibration was called dither from the Middle English verb 'didderen,' meaning to tremble.'" Ken C. Pohlmann, *Principles of Digital Audio*, Fourth Edition (New York: McGraw-Hill, 2000), 46. In the context of digital sound processing and quantization errors, John Watkinson describes that "dither was first recognized in connection with video quantizing in the 1950s, but the definitive treatment of dither and audio quantizing is generally considered to be that due to John Vanderkooy and Stanley Lipshitz, published in 1984." John R. Watkinson, "The History of Digital Audio," in *Magnetic Recording: The First 100 Years*, eds. Eric D. Daniel, C. Denis Mee, and Mark H. Clark (New York: The Institute of Electrical and Electronics Engineers, 1999), 112–113. In 2000, Lipshitz and Vanderkooy themselves, together with Robert Wannamaker, note that Gammaitoni's 1995 article had been the first to "directly acknowledge the correspondence" between dithering and the broader phenomenon of stochastic resonance." Robert A. Wannamaker, Stanley P. Lipshitz, and John Vanderkooy, "Stochastic Resonance as Dithering," *Physical Review E* 61, no. 1 (January 2000): 233. Gammaitoni writes in his article that examples of the stochastic resonance effect have been observed in fields as diverse as "neurobiology (e.g., neuron firing), natural events (e.g., avalanches), laser systems (e.g., laser threshold), complex systems (e.g., bifurcations), chemical systems (e.g., activation threshold), and political sciences (e.g., electoral schemes)." Gammaitoni, "Stochastic Resonance," 4691.

[69] Michel Serres, *Genesis*, trans. Geneviève James and James Nielson (Ann Arbor: University of Michigan Press, 1995), 132.

of "high definition" hard- and software that promises a definition well beyond the CD-standard. What is more, it has also seen the unexpected resurgence of vinyl records. Some listeners, it appears, find vinyl's sonic qualities, specific noises, and lower definition preferable to higher definition digital formats. Nothing could more clearly underscore the difference between technically verifiable definition and concepts of fidelity based on changing attitudes toward technical specifications, aesthetic preferences, and ideals of vanishing mediators.

Notwithstanding this preference for lower definition sound carriers, the discursive basis underlying technical innovations from the humble phonograph in 1877 to the reign of the CD in the 1980s has remained the idealist drive for absolute transparency. This persistent myth of perfect fidelity, which was already apparent in Tainter's work on the graphophone, is perhaps most poignantly expressed in Dolby's rhetoric of an "ideal audio device" promising a "remarkable clarity of reproduction" by imposing "no audible limitation on the signal." Here we espy the crucial difference between technical discourses on sound definition and the idealist myth for which perfect fidelity should be the ultimate goal of technological sound reproduction. Whereas the latter implies some kind of faithfulness to the "real" world outside the sphere of reproduction, an original signal that can be reproduced without any "audible limitation," those technical discourses merely establish measurable parameters. As such, they need not imply an intrinsic relation between originals and copies.

With regard to the persistent myth of perfect fidelity, the case of quantization errors and dither therefore provides the most appropriate cut-off point for this history of the noise of sound media. The culmination of more than a century's worth of continuous efforts to reduce these ever-present noises—from squeaking hand cranks and needle scratch to surface noise and tape hiss—digital technologies seemed to hold out the promise of unprecedented definition and reproductive clarity. Nevertheless, the fact that they also required the reintroduction of random noise shows that audiophile dreams of vanishing mediators consistently run into the idiosyncrasies and limitations of physical media, which keep presenting engineers and musicians with new and different noises and distortions. A more incisive analysis of technological sound reproduction therefore requires a different attitude toward the role of noise in sound recording; one based not on the myth of perfect fidelity, even if only implicitly, but rather on the understanding that noise and distortion are intrinsic and instrumental to the sounds produced and reproduced by technical media.

Unperturbed by technical discussions and audiophile ideals, musicians often embrace the many ways in which the noise of sound media affect the sounds that they record and reproduce. The background noise on Mark Hollis's solo record encloses the other sounds on the song much like a frame bordering a painting. In this intimate, almost claustrophobic listening experience, one imagines this music being recorded in a small, enclosed room. Although it would have been perfectly possible to prevent or reduce much of this noise, doing so would have drastically changed the music and the way we listen to it. Consider an even more striking example, the soundtrack to *Patience (After Sebald)*, a documentary about German writer W. G. Sebald's novel *The Rings of Saturn*. To accompany the documentary, the English musician The Caretaker turned a recording of Schubert's song cycle *Winterreise* from the 1920s into a series of highly evocative sonic miniatures.[70] In The Caretaker's music, the sounds of Schubert's songs become distant remnants of faint piano notes and muffled voices—fragmented memories of the original recording. Sounds that were intended as the main act when this performance of *Winterreise* was recorded now assume a secondary position, as the scratches, distortion, and surface noise that accumulated on the shellac disc on which they were pressed are brought to the fore and become a work of music. The signal-to-noise ratio is turned upside down as digitally manipulated noise and distortion become the primary compositional material, turning the original recording of Schubert's evocative, romantic songs into a sonic meditation on time, loss, and melancholia.

Since they take noise not as a disruption, nuisance, limitation, or problem, but as material for musical creation, these examples restate the question "How much noise is necessary?" Necessary, that is, to convey a message, set events in motion, or make an organism grow? To (re)produce a sound, a voice, a song, a piece of music? The next chapter continues along the path toward answering these questions by mounting an in-depth analysis of the operations of Dolby's analog noise-reduction systems and the principles underlying dithering. By further deconstructing the myth of perfect fidelity, which still underpins most discourses on technical sound media, we can begin to consider how a different, less idealistic take on the noise of sound media might reconfigure the relations among music, media, and listeners.

[70] *Patience (After Sebald)*, by The Caretaker, Bandcamp, self-pub, 2012.