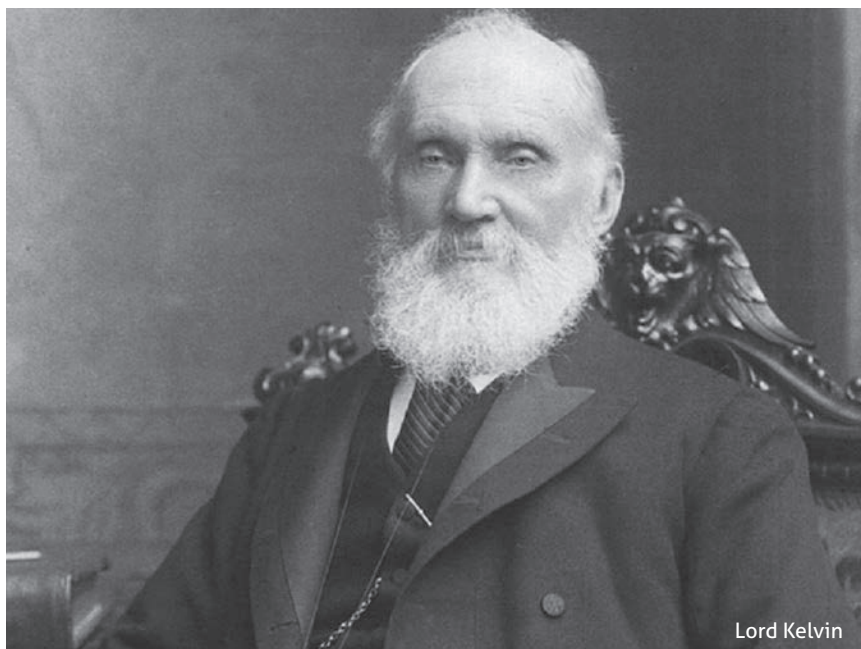


Let the Revolution Begin

MQA and the Overthrow of 20th Century Audio

Robert Harley

"There is nothing new to be discovered in physics now. All that remains is more and more precise measurement." —Lord Kelvin in 1900, five years before Einstein's paper on relativity.



Lord Kelvin

The term “paradigm shift” was coined by Thomas S. Kuhn in his influential 1962 book *The Structure of Scientific Revolutions*. According to Kuhn, a paradigm shift in the physical sciences occurs when a body of evidence accumulates that suggests the principles on which a scientific discipline is founded (the paradigm) is wrong, and a new paradigm replaces the old (the shift). Kuhn shows that this process unfolds in identical fashion throughout history no matter what the discipline.

Every scientific revolution, from the Copernican model supplanting Ptolemy's worldview, to relativity upending Newtonian physics, occurs in specific and defined phases. One of these phases is characterized by “crisis” in which a “battle” (Kuhn's terms) breaks out between followers of the old and new paradigms. The conflict arises because discoveries made within the existing paradigm don't quite fit that paradigm, yet the emerging paradigm is not yet accepted as scientific fact. Indeed, the emerging paradigm is often regarded as heresy.

Moreover, if you've spent your entire life adhering to a certain set of ideas, it's difficult to accept that your beliefs have been based on erroneous assumptions. You are simply too invested in the old paradigm. This resistance to new ideas is so entrenched that Kuhn suggests that a revolution is complete only when all the adherents of the old paradigm have died. He quotes Max Planck: “A new scientific truth does not

triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it.”

Andrew Quint's Guest Editorial in this issue prompted me to revisit *The Structure of Scientific Revolutions* (I first read it in 1990), because Andrew's description of the controversy over MQA mirrors Kuhn's “crisis” phase of a scientific revolution. As Andrew describes, some commentators have staked out the position that PCM (or DSD) encoding is essentially perfect and therefore MQA is unnecessary at best and a fraud at worst. Unfortunately, the Internet has given voice to anyone with a keyboard, allowing individuals with absolutely no understanding of MQA's technology, and no firsthand listening experience, to weigh in, often with vitriolic invective. There are even some respected experts in digital-audio technology and engineering who are skeptical of MQA.

These classic symptoms of Kuhn's “crisis” phase of a scientific revolution are the result of two distinctly different paradigm shifts on which MQA is based. Bob Stuart and Peter Craven (the British mathematician who co-developed MQA with Stuart) didn't invent the two emerging paradigms that are the foundations of MQA. Rather, they researched and discovered new ideas in other disciplines (specifically digital sampling in astronomy and medical imaging, and insights into psychoacoustics from neuroscientific advances) and applied those principles to audio. Other fields have been more open to these breakthroughs, but for some reason audio seems to be populated largely by calcified fundamentalists who cling to the past.

The first revolution MQA initiated (in the audio world, at least) is the idea that one pillar of digital audio, the so-called Nyquist-Shannon sampling theorem, while being correct for arbitrary communication, can be reconsidered for humans listening to music. Specifically, for “natural” signals such as sound or visual images (which tend to have

specific characteristics and statistics), limitations imposed by conventional sampling can be overcome by a more enlightened analysis and implemented with today's powerful digital-signal-processing technology. New techniques have been developed not for audio, but in other fields such as image processing and astronomy in which the resolution of closely spaced objects or the limitations of signal power can be of paramount importance. By exploiting signal statistics, cutting-edge medical-imaging technology can resolve data beyond the "Nyquist limit," allowing finer resolution of visual detail or flow. There's a direct parallel between resolving time information in musical signals and distinguishing between closely spaced objects in visual images. MQA has taken into account the human listener and the statistics of the audio signal, and adapted modern sampling theory to solve fundamental problems that have plagued digital audio since its inception. (For a technical primer on this topic, see my feature article in Issue 253 or on theabsolutesound.com. For an academic-level explanation, search on Google for "Sampling—50 Years After Shannon" by Michael Unser and "Sparse Sampling: Theory and Applications" by Pier Luigi Dragotti.)

The second paradigm shift on which MQA is based comes from neuroscience, specifically what the latest research has revealed about human hearing. Classical psychoacoustics (the old paradigm, originating in the 1930s) was based on experiments using test tones and beeps, with the subjects reporting on tones they could and couldn't hear. The researchers approached these experiments with two assumptions. The first was that the ear was a linear microphone-like device and the brain a passive receiver that analyzed the electrical impulses, creating the sensation of sound. The second assumption was that the brain was a frequency analyzer, and that our auditory system's resolution of timing information was implicit in our upper-frequency limit (as dictated by Fourier analysis). That is, we couldn't discriminate timing information that implied a bandwidth greater than 20kHz. Consequently, psychoacoustics until very recently was primarily focused on the frequency domain: which pitches humans could hear, at what thresholds, along with related phenomena such as masking and the concept of critical bands.

This primacy of frequency, and the belief that the ear was a passive pick-up, has informed and permeated audio engineering ever since Harvey Fletcher's famous experiments at Bell Labs. This paradigm, while useful in many ways, unfortunately led us astray. Audio engineering has since its birth revolved around frequency-related criteria because it was simply reflecting the psychoacoustic paradigm, leading to the design principles, instrumentation, and analysis tools used to this day. The limitations of this primitive model of how humans hear reached its grotesque zenith (or nadir, if you prefer) in MP3, which is theoretically perfect according to the old paradigm of the ear as a passive linear receiver and the brain as a frequency analyzer. In the early 1990s I attended Audio Engineering Society conventions in which Karlheinz Brandenburg, the lead developer of MP3, presented papers describing his research. His hubris was on full display, as he casually used terms such as "psychoacoustic redundancy" and "informational irrelevance" to explain how throwing away 90% of the bits was a good thing. We all know how that turned out (except for the Fraunhofer Institute that supported the research, which at one time reaped \$100 million per year in MP3 royalties). But in Brandenburg's defense, he was operating within the old psychoacoustic paradigm developed fifty years earlier.

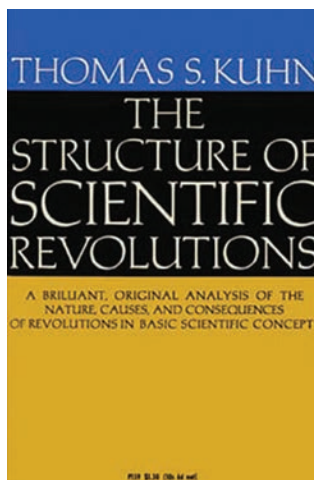
The new psychoacoustic paradigm recognizes that human hearing didn't evolve to hear tones and beeps. Rather, it is exquisitely tuned to detect the sounds of nature, which aren't composed of frequencies and tones, but of transients of indeterminate and randomly varying frequency. In fact, many of the sounds of

The researchers at the cutting edge acknowledge that we're still in the infancy of understanding how the neural pathways operate.

the natural world, an understanding of which confers important survival benefits, have *no* frequencies. Examples include crackling leaves, snapping twigs, the sounds of wind, rain, and running water. Our hearing system is highly adapted to operate against this background of natural, causal sounds. Not surprisingly, therefore, the latest neuroscientific research reveals that our hearing mechanism is considerably more dependent on, and sensitive to, timing cues than on frequencies. Moreover, this research has produced startling revelations that would be far beyond the ken of the early psychoacoustic

researchers. For example, there are more neural pathways descending *from* the brain to the ear than from the ear to the brain. Why would our hearing system benefit from this two-way communication? Modern neuroscience and modeling reveals that the brain is constantly sending signals to the ear, modifying its response in real time, as we are perceiving the sound. As we listen, signals from the brain physically "tune" the ear to better encode the specific information it needs to more accurately determine exactly *what* is creating the sound and *where* it is coming from. These

signals descending from the brain adjust both the cochlea and the ascending neural pathway, fine-tuning the auditory system's so-called "grouping" and "feature extraction" abilities. The ear and brainstem response is constantly changing microsecond by microsecond. (This phenomenon, incidentally, is one reason why lossy codecs such as MP3 fail in practice despite working in theory; the masking model on which they are based views the ear as a passive device. It's not nice to fool Mother Nature.) The implications of this discovery cannot



be overstated. The fact that neurons change their coding in real-time to combine and extract features in the sound tells us that the system is highly non-linear. It also suggests that simplistic theories based on Fourier analysis—and the closely related sinc (cardinal sine function) sampling kernel on which digital audio’s low-pass “brickwall” filtering is based—must be viewed with caution.

Digital audio systems introduce specific errors, correlated with the signal, that smear sonic events in ways that never occur in nature, confusing the brain and reducing its ability to recognize and identify those sonic events. In the last decade, researchers have independently confirmed that we are much more sensitive to timing information and temporal microstructure than predicted by our 20kHz frequency limit. (See, for example, “Human Time-Frequency Acuity Beats the Fourier Uncertainty Principle” by J.M. Oppenheim and M.O. Magnasco, *Physical Review Letters*, 2013.) It’s this timing information that helps the brain perform the apparent miracle of converting neural impulses into the impression of hearing individual objects in space. Degrade that information, as digital audio encoding and decoding does, and you reduce the clarity with which we perceive objects in natural space. The recent psychoacoustic research also reveals that one group of neural pathways from the brain to the ear is dedicated to transmitting only reverberation information, which is a critical part of the natural world (and also of musical realism). So although we can’t hear test tones above 20kHz, we *can* detect the benefit of temporal microstructure within midrange frequencies right down to the microsecond level. The list of radical new discoveries goes on and on, revealing that our hearing mechanism is exquisitely more complex and sophisticated than previously believed. Yet the researchers at the cutting edge acknowledge that we’re still in the infancy of understanding how the neural pathways operate.

Stuart and Craven have studied this academic research and applied it to understanding the different ways in which distortions of temporal microstructure affect our ability to identify, segregate, and locate “external” objects—to experience a well-defined soundstage, in other words. To quote Stuart, “The more we stop interfering with the microstructure, with the stop and start of sonic elements, the easier it is for our brainstem to stream the necessary components for the perception of the viola, of the violin, of the piano, and the easier it is to ‘grasp’ the sound of the venue before the first note is played.” Indeed, in my comparisons of the same music in conventional digital and MQA (made from the same master), I can often instantly identify the MQA version as soon as I hear the room sound by its more realistic sense of space.

So, here we are in 2017, with our digital-audio systems designed around first-generation paradigms of information theory (Nyquist-Shannon) and psychoacoustics (frequency-based, the ear as a linear and static device). MQA comes along and forges a new path, building on the advances in other fields and developing from first principles an entirely new way of looking at the question of how best to encode, distribute, and decode digitally represented music. By focusing on the entire analog-to-analog chain, the result is a system that delivers sound quality better than that of the original high-bit-rate master recording (through correcting technical errors in the original A/D converter); is backward-compatible with all distribution platforms and playback hardware; offers assurance (via the MQA light on every DAC) that the bitstream being decoded by your DAC is identical to the bitstream created in the studio; and creates a file that is small enough to be streamed to everyone. It’s quite astounding that MQA can combine so many virtues, and solve so many problems, in a single stroke. It’s an audiophile’s dream come true.

Every scientific revolution begins when discoveries are made that aren’t explained by the existing paradigm. To cite one example of this in digital audio, high sample rates sound better than lower sample rates, even though the upper limit of human hearing is regarded as 20kHz. According to Nyquist-Shannon, the CD’s 44.1kHz sample rate can perfectly reconstruct the audio waveform all the way up to 20kHz. And according to first-paradigm psychoacoustics, information above 20kHz is irrelevant, and our temporal resolution is limited to that implied by that 20kHz upper-frequency limit. If this is the case, why would higher-sampling rates sound “better?” The answer is that



the digital filters required by Nyquist-Shannon sampling at 44.1kHz introduce time distortion, or “temporal blur.” The filters for higher sampling rates are gentler and thus introduce less temporal blur. (Specifically, CD introduces around 5ms of temporal blur; 192kHz/24-bit PCM creates 300μs of blur; MQA aims for end-to-end analog blur as low as 10μs; MQA actually targets a system response similar to that of sound traveling a short distance in air.)

This, and many other anomalies that didn’t fit the existing paradigm of PCM digital audio theory and psychoacoustics have led us to Kuhn’s “crisis” phase of the revolution. The existing paradigms are showing their weaknesses, and new paradigms are emerging in which these anomalies are no longer anomalies, but fully consistent with, and explained by, the new paradigm. MQA is thus in the crosshairs of Kuhn’s “battle” between those who cling to the old paradigm and others who embrace the new. Early PCM audio (and DSD) will one day be regarded as primitive relics of the past, the product of first-paradigm thinking in audio engineering, information theory, and psychoacoustics. But as Kuhn demonstrates with example after example, it will be a long time before this revolution is fully complete.

Viewed in the context of Kuhn, it’s not surprising that MQA has its critics. MQA fits the definition of a paradigm shift; the ideas on which it is based are not advances within an existing framework of knowledge, but represent an entirely new framework. It’s the new framework that some people can’t comprehend, along with a reluctance to abandon long-held beliefs in certain “proven scientific facts.” But I suppose we should cut the critics some slack. After all, if Lord Kelvin could have been so wrong about the state of physics in 1900, it is easy to understand how a few audiophiles could be so mistaken about MQA. **tas**