

**TECHNOLOGY AND THE DEGRADATION
OF MUSICAL AESTHETIC EXPERIENCE:**

**OBSERVATIONS ON THE HISTORICAL
CORRELATION BETWEEN THE
INCREASING TECHNOLOGY
AND DECLINING QUALITY
OF MUSICAL RECORDINGS**

Jerald Hughes

#CIS-2002-06

May 2002

Jerald Hughes
Doctoral Student
Department of Statistics & Computer Information Systems
Zicklin School of Business, Baruch College
Jerald_Hughes@baruch.cuny.edu
cisnet.baruch.cuny.edu/phd/hughes

© 2002

The CIS Working Paper Series provides for early dissemination of research produced by members and affiliates of the Computer Information Systems program at the Zicklin School of Business, Baruch College, City University of New York and the Information Systems program of study of the PhD in Business program of the City University of New York.

Please address all inquiries to:

CIS Working Paper Series
Zicklin School of Business
Baruch College, Box B-11-220
17 Lexington Avenue
New York NY 10010

**TECHNOLOGY AND THE DEGRADATION OF
MUSICAL AESTHETIC EXPERIENCE:
Observations on the Historical Correlation Between the
Increasing Technology and Declining Quality of Musical Recordings**

ABSTRACT

After reaching a high point in consumer quality in the 1/4" tape analog format, musical recordings have undergone a continuous and serious decline as newer formats have been introduced. Ever since the early 1970s, issues of production cost, ease of use, and widespread distribution have taken precedence over the quality of the actual musical content. In recent years, largely as a result of the Internet phenomenon, the trend has accelerated, driving the fidelity of musical recordings down precipitously. The 8-track tape and tape cassette formats were the first to make the tradeoff of convenience for quality. The audio CD format did away with surface noise problems, but has fallen short in other areas. Internet formats such as MP3 and streaming audio have accommodated bandwidth limitations by providing much lower quality recordings. Further technological improvements will eventually be able to bring the quality of recordings up to an acceptable standard. Future applications of the computer to artwork will have their strongest points not in the reproduction of older art forms, but in the creation of wholly new genres of artworks.

**TECHNOLOGY AND THE DEGRADATION OF
MUSICAL AESTHETIC EXPERIENCE:
Observations on the Historical Correlation Between the
Increasing Technology and Declining Quality of Musical Recordings**

TABLE OF CONTENTS

I.	Statement of Theses.....	1
II.	Introduction.....	2
III.	Analog Formats.....	4
IV.	Audio CD Format.....	9
V.	Internet Formats – MP3.....	14
VI.	Internet Formats – Streaming Audio.....	18
VII.	Future Developments.....	19
VIII.	Computer Art.....	21
	Appendix A – Endnotes.....	24

STATEMENT OF THESES

THESIS: The more we apply technology to musical aesthetic objects, the lower the quality of those objects, compared both to originals and to prior analog and in some cases even prior digital reproductions. What is communicated in these objects is nothing more than distortions of the raw facts about the objects' form. Their essential and critical content, which is aesthetic experience, is lost.

COROLLARY THESIS: Technology is making possible new types of aesthetic experiences, which were not possible with traditional art objects. The most mature of these at present is “immersion”, by which is here meant the use of combinations of multimedia

tools and creations to create a global, holistic sense of 1) a created reality and/or 2) the mind of the author.

INTRODUCTION

The widespread access of consumers today to devices which can store, process and reproduce digital images and sounds has resulted in an explosion in both the production and distribution of such multimedia objects. On audio CD's, CD-ROMS, DVD disks, digital cameras, and the Internet, realistic (relatively speaking) sounds and images are available everywhere. What has been little discussed amid this flood of on-demand sensory experience is what the costs of this transformation have been. In moving from the real to the digital, have we unwittingly made trade-offs whose impact we did not anticipate? This paper will argue that for musical aesthetic objects, we have indeed, and that what has been lost is considerable.

The essential and irreducible justification for all these new devices and modes of display is content. Somewhere, perhaps publicly accessible and perhaps not, is an original drawing, painting, photograph, film, studio tape recording, or even live theater or concert performance. Computer graphics notwithstanding, the vast majority of objects which exist for the sake of their own form and content, and are distributed and viewed for the purpose of experiencing those qualities, are still in the realm of the tangible, the hand-made. This foundational level of presentation provides the richest possible aesthetic experience. Seeing and/or hearing a work of art at first hand is an experience that cannot be even roughly approached by any form of reproduction. The physical presence of the work itself can be overwhelming, as in Lichtenstein's *F-111*, Dali's jewel-like *The Persistence of*

Memory, Rodin's The Burghers of Calais, or Vermeer's mesmerizing The Lacemaker. Hearing Mahler's *Adagio* performed by L'Orchestre de Paris, or Cavaradossi's opera *I Pagliacci* as sung by John Vickers in the concert hall is a wrenching, visceral, physical experience, in the impact of the sounds made by dozens of instruments played by virtuosi, an impact that itself cannot begin to be duplicated by even the most superior electronic sound system. The stereo system may pass along approximations of pitches, timbres, and rhythms, but cannot recreate the living presence of the music itself, as it was originally intended for live performance. The same is true, by the way, even with performances containing far simpler musical and dynamic content, such as those by, say, the rock band Led Zeppelin or blues artist B.B. King. When we encounter the original work of art in the flesh, so to speak, we have come as close as it is possible to come to the *ding an sich*, the "thing in itself". The artist's hands have touched the objects we observe, shaped and molded them, and thus we are at only one remove from the artist himself. Live theater and live musical performance bring us even closer, since the artwork itself is ephemeral; it exists only for those moments during which it is performed, and its creators are physically before us in the form of actors, musicians, and conductors. Though the artifice of the theater, with its budgeted sets and costumes, artificial lighting, and the unavoidable "fourth wall", is brought continually to mind, the living presence of the performers themselves more than compensates. Indeed, the emotional component of the experience is multiplied by the fact that fallible humans are performing live in front of you. These things are the gold standard. They are the original form of the content, and the sole justification, the *raison d'etre*, for the existence of all of the billions of reproductions now at hand. The aesthetic experiences we derive from our encounters with the originals are the purest and

the most intense possible. All of the discussion which follows will refer to this level as a baseline—necessarily, so that we can keep in mind what we have lost when we translate art objects into other media.

[In an important sense, the preceding paragraph suffers from the same problem discussed throughout the paper. I am attempting to capture in words alone what cannot be expressed even by the highest quality reproductions of images, sounds and words: the essence of real aesthetic experience, which is the direct encounter with the artwork itself. For those who have not been fortunate enough to experience such encounters with the artworks mentioned above, I can at best only refer the reader to his or her own first-hand experiences, and recommend some blocks of future time set aside for attendance at plays and concerts, and a few visits to galleries and museums. If the reader has not yet grasped the value of the quality referred to above as the baseline, no amount of argument and explanation here will suffice. By its very nature, that quality must be grasped at first hand; indeed, that is the main point of the argument being made.]

ANALOG FORMATS

The earliest instance of artworks being made broadly available in digital form occurred in the late 70's, in the shift from analog vinyl LP's to audio Compact Disks. It makes sense that the changeover would have occurred here first, since the amounts of data involved, great as they are, are far smaller for sound than for any of the visual forms. The reason is that we experience sound as a simple linear stream (actually, two simultaneous streams, but the point is that the form of the data is one-dimensional, whereas

reproductions of visual data start at two dimensions at a minimum, so that the increase in amounts of data for visual reproductions is geometric). The basic physical facts involve nothing more than smoothly and swiftly changing levels of pressure against the eardrums. Storing and reproducing the sounds thus means finding ways of encoding those levels. The reference standard for analog encoding was, and still is, studio tape recording. That is, fluctuating air pressure levels are converted to electrical signals which vary in voltage in precisely the same way as the air pressure; those electric signals are then preserved on tape as domains of magnetic alignment on tape media. The microphone accomplishes the first step: a portion of the microphone known as the capsule vibrates (and therefore moves) in sympathy with the vibrations in the air which strike it, and those motions are converted to voltages. The second step, preserving those electrical signals, employs a tape base which is coated with a ferric material, usually some proprietary formula using iron oxide as the primary component. The ferric coating should be very fine-grained, so that it can respond quickly and with high resolution to subtle changes in the signal, and very dense, so that as great a separation as possible between the lowest and the highest signal strengths can be attained. All of these factors, we will see later, have their counterparts in the digital realm. In the hands of an experienced engineer with very sensitive (and very expensive) equipment, the studio tape can provide an extremely high-quality aesthetic experience. In fact, the process of analog recording to tape itself actually adds certain qualities (as natural physical effects of the equipment being used) produced in the form of content-dependent harmonics (basically, integer multiples of the source frequencies, although the interactions get much more complicated than this) which actually enrich and enhance the sound; these qualities do not appear at all in digital recordings, and their passing was protested, then

mourned, by audiophiles who opposed (and in many cases still oppose) audio CD's and digital recording in general. For awhile in the late 60's and the 70's, this level of reproduction was made available to consumers at large in the form of reel-to-reel tape decks playing 1/4-inch tape; however, these devices were expensive and cumbersome, and were adopted only by the most dedicated audiophiles.

Up to this point, the history of audio reproduction had been one of continual improvements in fidelity. From wax cylinders to vinyl 78 rpm records to the 33 1/3 LP to the stereo LP album, from wire recording to the sophisticated 24-track 2-inch studio tape, the march of technology had been successfully put in the service of improving the aesthetic experience. Indeed, the point had been reached at which even music reproduced on home stereo equipment could have profound effects; William Styron writes of a sudden and dramatic turning point in his years-long struggle with depression, which occurred as he listened to a broadcast of the Brahms *Requiem*, an extended and complex work for full orchestra, soloists and chorus¹. Unfortunately, the next application of technology to sound, still in the analog realm, would be a giant step backwards.

The equipment needed for either vinyl LP's or reel-to-reel tape is large, heavy, and for all intents and purposes non-portable. The invention of the transistor radio had given consumers an expectation of being able to take their music with them wherever they went; what was needed was a distribution vehicle that allowed consumers to take their own recordings with them. The first attempt, which achieved broad success for awhile, was the 8-track tape. This format preserved the actual size of the tape (1/4" width), although size is deceiving, because the tape was divided into 8 separate sectors, 4 sets of stereo tracks. Thus, far fewer magnetic particles were devoted to preserving the electro-magnetic signal,

and the result was a lower signal/noise ratio. Another handicap of the format was the fact that the musical program had to be cut into 4 separate segments, with unavoidable gaps between each as the player shifted from one pair of tracks to the next. In effect, consumers were making a tradeoff between portability and quality; this is the essential shift that still lies behind the ongoing decline in digital aesthetic objects today.

The next step backwards was also the last important innovation in the analog realm: the cassette tape. The miniaturization of the recording and playback head components, and improvements in manufacturing processes which made possible fairly precise alignment of those heads with the tape, made possible a much smaller tape, which could hold an entire side of an LP album's worth (30 minutes max.) of recorded sound without breaks. The early tape formulations from the late 60's were of very poor quality compared to those available today, and the cassette tape was not considered a viable distribution vehicle by those accustomed to LP's until some years later. The cassette tape's severe physical limitations meant that even fewer magnetic particles were used to preserve the signal; the result was that the signal was poor to begin with, and was vulnerable to even tiny errors. The slightest misalignment of the heads basically destroyed the signal, and the loss of magnetic particles as the ferric coating flaked off the tape caused a severe type of distortion known as 'dropout'. Gradual improvements in the quality and durability of tape formulations have mitigated these problems somewhat, but even the best standard consumer music distribution cassettes have a Signal/Noise ratio of only 52 db, and a frequency response that tails off steeply above 13 kHz² (human hearing extends up to about 20 kHz). Compare this to the specifications of the much older reel-to-reel tape technology: S/N ratio of 75 db, and frequency response up to 15 kHz³. These are

immediately audible differences in quality, even for casual listeners. Fine-tuning the high-quality decks used in recording studios can extend performance considerably, with 90 db of S/N ratio and frequency response up to 35 kHz⁴. This kind of ‘hot-rodding’ is generally done by modifying mechanical components to increase the transport speed from a maximum of 15 inches/sec to 30. Common sense tells us that this makes far more magnetic particles available to preserve the signal strength (S/N ratio), and the speed means that finer-grained changes can be detected and stored (35 kHz ceiling).

The cassette tape is a very poor distant cousin to the reel-to-reel consumer decks which preceded it. The noise floor was so high that consumers complained about the disturbing level of hiss. The music distribution industry responded with an ingenious solution, Dolby Noise Reduction. This method artificially boosts the high frequencies of the source program material before it is printed to tape (a procedure known as Dolby encoding). The Dolby-equipped machine then globally reduces the most troublesome frequencies involved in tape hiss; the result is that the frequency levels of the program material are reduced to their original intended balance, while the tape hiss noise floor is noticeably reduced, improving the S/N ratio by about 5 db. Unfortunately, the only reason Dolby can make such a large improvement is that cassettes are so bad to begin with. The best tape formulations (not used by record companies on tapes for distribution of music to consumers) used on the best Dolby equipment can only do as well as 64.5 db S/N ratio, with close to a 14 kHz frequency-response ceiling⁵. The result was that cassettes were never accepted as a medium for high-quality sound reproduction, but did achieve widespread use in portable players, especially after the SONY Walkman was introduced. The recording industry came away from this experience with a critical conclusion:

consumers will accept lower sonic quality in exchange for portability and ease of use. This conclusion has colored every subsequent development in the field of music distribution vehicles for consumers. Later on, however, the argument will be made that this tendency has been pushed too far, and that consumer response to the very latest technological trade-offs proves that there is a limit to this strategy.

AUDIO CD FORMAT

For most consumers, the first encounter with purely digital aesthetic objects came with the introduction of the audio CD. Several considerations determined the specifications of this format. The disc had to be able to store the same amount of music program material that consumers were accustomed to buying on vinyl LP's and cassette tapes, about 60 minutes maximum. The disc had to be playable on low-priced consumer equipment, so the specifications for the precision of the laser and the size of the reflective pits on the discs were set with generous margins for error. These two factors in combination, along with a preference for small size (large video laser disks were adopted by very few consumers), allowed for the CD-quality audio settings still in use today: 16 bit sample depth, at 44.1 kHz sample rate. The latter reflected a straightforward (and flawed, it turns out) application of the Nyquist Theorem. Humans, under ideal conditions, can hear isolated frequencies up to about 20 kHz (although a few gifted and well-protected ears can hear significantly higher). The Nyquist Theorem says that in order to be able to capture this frequency with digital sampling, we need to measure the voltage level of the signal at least twice during one cycle of the waveform. Thus, the sample rate chosen was roughly twice the frequency ceiling of human hearing. Digital sound storage did confer one real and serious benefit,

which was the complete eradication of surface noise. Imperfections in the surface of the vinyl disk, along with the accumulation of scratches and dust, had always been a serious impediment to high quality for the LP format. Since the numbers on the CD audio disk were the source of the output waveform, the digital medium itself was totally transparent. That is, the quality of the sound achieved in the digital recording studio could be transported, without any diminution in quality whatsoever, to the consumer's playback device—at which point the sound quality finally would diminish significantly, since consumer playback devices, amplifiers, and speakers would all be of lower quality than those in the studio. Vinyl LP's, when produced at very high quality and handled with great care, could equal this level of performance⁶, but most consumers did not bother to take the steps necessary to preserve their records, and only a few record companies invested in the level of production equipment and materials quality required (Deutsche Gramophon was a notable exception).

Although some segments of the consumer market resisted the audio CD, the audiophiles because of flaws in the sound, others because of the huge price increases applied to the new format (prices nearly doubled, from \$6 to over \$10 per album), others because of the initially very high costs of playback equipment, the audio CD eventually achieved complete penetration of the consumer music market. The advantages were those of convenience: small size, high portability, high durability and resistance to damage compared to vinyl LP's. The disadvantages, relatively subtle ones, were those related to the aesthetic experience.

Both the sample rate and the bit depth chosen for the audio CD are problematic. Audiophiles immediately complained that the new format sounded brittle, cold and edgy.

Although such comments were initially dismissed by recording industry experts as format-snobbishness, it turns out that the criticisms are justified in many musical contexts.

The sample rate determines the frequency response. Although the Nyquist Theorem was thought to have guaranteed satisfactory performance for human ears, it turns out that there are several factors in acoustics and human hearing that were not taken into account. One is that natural sound sources, such as musical pipes or guitar, piano, and violin strings, create not single pitches, but instead pitch complexes. The simplest analysis includes consideration of the integer multiples of the basic frequency, called the fundamental. At least the first eight overtones, as they are called, are significant in our perception of the timbre of a particular sound. When we combine multiple pitch sources, such as 6 piano keys struck at the same time, we're already dealing with at least 54 different pitches (6 fundamentals plus 8 overtones apiece). In addition to this, all of those pitches interact with each other to form 'difference tones', which are additional perceptible pitches whose frequency is the difference between two source frequencies. There is an electronic musical instrument based upon this principle alone, the Theremin. The Theremin has an ultrasonic fixed 'carrier' frequency somewhere in the range of 175,000 – 225,000 Hz, a range far beyond the limit of human hearing. Another frequency is also produced in the same range which can be modulated, so that the relationship (difference) between the two can vary from 0 to about 2000 Hz. Although humans don't hear these two frequencies, we do hear the Theremin's 'difference tones' loud and clear⁷. The implications for digital audio are clear: if we limit the frequency response of the medium through our choice of sampling rate, then an entire universe of overtones and difference tones will be lost. Although the casual listener is unlikely to be consciously aware of the lack of secondary frequency

material when listening to an audio CD in isolation, many listeners will be subconsciously aware of a difference they cannot describe, and most will notice immediately the lack of richness in CD program material when it is compared side-by-side with live performance or high-quality reel-to-reel tape. To put it another way, the natural phenomenon of “difference tones” brings normally inaudible very-high-frequency program content down into the range where humans can perceive it. The audio CD sample rate cannot capture any of this high frequency program material, so all of the rich sonic qualities contributed by the difference tones from this range are simply absent.

Sample rate also comes into play when we consider the importance of phase relationships (on multi-driver speaker systems) and stereo imaging. It turns out that the human hearing system can discriminate signal arrival times at the two ears with differences as low as 5 microseconds, which is far shorter than the time difference between two samples at 44.1 kHz. Many audiophiles describe digital CD’s as having a ‘flattening’ effect on the program material; timing problems caused by lack of resolution in the sample rate may be behind this criticism. Resolution in the analog media, by comparison was basically determined by the fineness of grain of the iron oxide molecules on tape, and of the vinyl molecules on LP. In order to preserve these timing relationships, we would need a sample rate of 200 kHz⁸.

The bit depth also has problems. The evolution of the encoding of sound in digital form has proceeded by bytes. The first experiments used just one byte, 8 bits. This gave only 2 to the 8th power different possible signal levels, a very rough approximation of a waveform from any source. Early PC sound cards, with their severe main bus, clock speed and storage limitations, offered only 8-bit sound. Although audio recorded this way was

minimally adequate for simple applications like signaling computer functions and errors, it was nowhere near good enough for music distribution companies. When a continuous waveform is digitally recorded, the resultant data file consists of a sequence of discrete numbers, essentially a step function that approximates the original waveform. It is the job of the Digital/Analog converters in the playback devices to change that sequences of vertical steps back into a smooth wave function—the circuits and algorithms used for this purpose vary widely in both quality and price, with consumer devices, as usual, occupying the low end. If the steps in the digital data are large—and the steps in the 8-bit format are very large—then distortion is produced called quantization noise, which we hear as hiss, white noise. This noise, unlike that present in analog circuits, which is thermal noise, cannot be removed or even reduced by any means whatsoever. It is inherent in the format. With only 8 bits to encode audio, the S/N ratio is just 40 db—worse than even the cheapest cassettes⁹.

The solution for audio CD's was to add an extra byte, yielding 16 bits for storing each sample number. This resulted in a vast improvement in the S/N ratio, all the way up to 90 db¹⁰; the best vinyl LP's had managed only about 65 db, with average systems being much lower, in the range of 45-50 db¹¹. This solved the problem for some types of program material, but not all. As long as the source material fills a relatively narrow range, as is common for pop music, 16 bits is adequate. For programs which vary greatly in loudness and softness, which includes most classical music and much jazz, quantization noise is still a significant problem. The top digital signal level is an absolute ceiling for the encoding of audio data; once that level has been set, everything else must be encoded in relation to it. This means that programs which include both very loud and soft passages

will be encoded with too few signal levels in the softer passages. Especially the reverb tails of live recordings can be heard, even on commercial releases, breaking up into a buzzy, grainy distortion.

When all the factors are considered, the CD giveth, and taketh away. For the listener concerned with the aesthetic experience, and leaving convenience and cost considerations aside, the CD helps by doing away with the LP pops and clicks. On the other hand, the quality of the sound of the program material itself suffers; in the case of sensitive source material, it suffers dramatically.

If the introduction of the audio CD was a mixed bag of aesthetic considerations, the next two technological innovations, MP3's and streaming audio, are unmitigated disasters.

INTERNET FORMATS – MP3

Between roughly 1990 and 2000, personal computers passed several critical thresholds which allowed them for the first time to usefully store and process complex sounds and images. Although mainframe computers had been able to deal with images and sounds in sophisticated ways since at least the 70's—an example would be the IBM 370 machines used at NASA's Johnson Space Center (Mission Control in Houston) to produce and display high-quality flight simulations for use in training space shuttle pilots—these applications were enormously expensive and therefore experienced only by very few persons. Microcomputers were also capable of highly rudimentary graphics and sound from the very early days, as far back at least as 1978, when machines like the Commodore, the Amiga, and the Apple II were capable of producing a few blocks of color and pitched sequences of beeps; however, these images and sounds were so crude that they were

severely limited in use, and bore no comparison at all to the images and sounds we as a culture were accustomed to in the analog realm: live and filmed drama, musical concerts and recordings, paintings and photographs displayed in galleries and museums, and reproduced via four-color processes on paper.

Reproducing these required irreducible minimums in the areas of central processor speed, main bus speed, disk storage capacity, screen resolution, color and sound bit depth, and RAM capacity, among others. For most of the history of the PC, high-quality speakers were not widely available; nor were they desired, since there was nothing to playback on computers that required high quality. The result was that, even when the minimum thresholds for computer audio were achieved—386 machines could do it, with difficulty, while 486 machines and HD storage in the hundreds of megabytes made serious computer audio possible for the first time—no one used them for that purpose. The driving force that made PC's into expensive MP3 players for millions was the explosion of Internet access. [Since we're talking about issues of audio quality, the discussion which follows should be assumed to apply only to the highest quality MP3 encoding settings, which are variable.]

MP3 (MPEG 1 Layer 3) encoding, created by the Fraunhofer Institute¹², exploits a psychoacoustic principle known as masking. The theory is that our human ears do not hear everything that is present in a complex waveform. Loud sounds prevent us from hearing simultaneous softer sounds. If you can't hear what your friend is saying to you on the subway platform as the train thunders by on the express track, this is why. Soft sounds which follow loud sounds are also masked, as the nervous system recovers from the strong signal. The MP3 separates the source material into hundreds of frequency bands and throws away those which the masking algorithm determines will be largely inaudible. The

result is a data file which is small enough to be distributed over the Internet (with difficulty over dial-up modems; with ease over broadband connections such as those available at universities). The propaganda associated with MP3 files and their players refers universally to “CD quality”—as though this were a virtue, instead of a flaw, as we have already seen. Despite such claims, actual tests show that the compression algorithm has made a compromise between file size and audio quality, one that is perceptible for even casual listeners¹³. As with the audio CD format, the performance of MP3 varies with the type of source material involved. Program material which includes very low signal levels should never be encoded with MP3, since the algorithms will throw away such material, leaving behind nothing at all.

MP3, unlike the other audio media we’ve looked at so far, is a compressed format, so it is not directly comparable in terms of the performance measures we’ve been using. The de-compressed MP3 file produces a stream of samples comparable in sample rate and bit depth to audio CD’s, so quantization noise is kept at low levels. On the other hand, the compression itself introduces significant distortion artifacts. Both the research into masking effects and its applications in MP3 are imperfect, and achieving useful levels of compression involves some compromises. The results include loss of high-frequency response, loss of stereo imaging, and poor transient response, among others. Overall, the effect is roughly equivalent to a cheap cassette tape. Since most MP3’s are listened to on cheap headphones and low-quality computer speakers, the loss of quality is not much of a consideration, since the playback hardware would do little better even with high-quality source material. The only advantage of the MP3 is convenience: the files are small and easy to obtain, and with today’s large hard drives, the expense of storage is basically zero.

Since they are pure information, the physical storage space is also zero, unlike audio CD's. The disadvantage is that, by choosing MP3's the user is returning to the audio quality of cassette tapes of about 1970.

Internet music distribution schemes have met severe resistance from users, no matter what the payment model. Although this is usually attributed to the notion that MP3 users are unwilling to pay for music they've already become accustomed to getting for free, the fact remains that the demand for MP3 files does not include the willingness to pay even very small fees for them. Recall that consumers would not pay for popular music distributed on cassettes until the audio quality of that medium reached a minimum acceptable level. It may be that the same phenomenon is behind the failure of MP3's to generate profits as a music distribution medium. Regardless of the hype, most music consumers still have at least a subconscious knowledge of what a sonically rich musical performance entails, and they are unwilling to commit money to a medium which is too distant from that reference standard.

The only reason to have MP3's in the first place is so that users can obtain the music files via direct downloads over the Internet. And the only reason draconian compression schemes are required is to overcome current bandwidth limitations. Thus, in the long run, the use of any compression algorithms at all will be beside the point. Once it becomes the norm for individual residences to have fiber-optic links at the local-loop level, downloading the entire contents of a 700-MB CD or even a 4.3 GB DVD will be passe. That day is still a long way off, so in the meantime we have even worse schemes to deal with, like streaming audio.

INTERNET FORMATS – STREAMING AUDIO

Streaming multimedia content on the Internet, by which is meant real-time delivery, without waiting for significant downloads (there is still a short wait while the initial buffer is established), is still at the nascent stage of development. The current state of the art is like the Great Dane that can walk on its hind legs: it's not that the thing is done well, we are simply amazed that it can be done at all. The "quality", if that term is even appropriate in this context, does not bear even a moment's comparison with any other audio format available today—streaming audio is truly bad. Frequency response which often tops out below 8 kHz, myriad distortion artifacts from the radical compression ratios employed (often in excess of 100:1), dropout from lost Internet packets...the streaming audio format is so far removed from the original source material that it may be impossible to even recognize the performer's voice, hear the instruments involved (bass is virtually non-existent in streamed audio), or even hear the chords. For the sake of mere convenience, instant access via the Internet, we have given up virtually everything that we prize about music in the first place.

And yet streaming audio has found a niche: the popularity of Internet "radio", by which is meant live programs of music, news and talk sent out over the Internet, has exploded. Thousands of "channels" are available, many of which specialize in restricted genres of music which gets no airplay on commercial or even public radio. Universal access to rarely heard musical forms is not an insignificant benefit; indeed, it exemplifies the best of the fundamental societal benefit that the Internet confers: access to information. At present, however, the beneficial effect is deeply undermined by the execrable quality of

the information product itself. No one who has ever observed, say, a live flamenco performance, will admit any comparison at all between the live performance and the streamed audio file. What is lost in the translation is just barely short of everything; only bits of melody and occasional rhythmic accents remain. The fire and excitement that are the essence of flamenco are entirely absent from streamed audio versions.

This brings us to the present day. What comes next is both better and worse.

FUTURE DEVELOPMENTS

Good news first: the audio standard is eventually going to migrate to a 3-byte sample depth (24 bits). This degree of resolution will meet and surpass the limits of human perception; in other words, 24-bit audio will be able to produce sound files of such high quality that they will extend significantly beyond the physical limitations of human hearing. Equipment capable of recording at this bit depth is already available, and has made significant improvements in the results heard on audio CD's. Music distribution at present is really a hybrid system. The source material is recorded at 24 bits, but must be reduced to 16 bits for production on standard audio CD's. This process can raise problems of its own, especially if the sample word is merely truncated, but effective sound-shaping algorithms known as "dithering" are able to make the transition more or less graceful. Sample rates are another story. Some recording systems are capable of a 96 kHz sampling rate, but as we saw in our discussion earlier, this may not be enough. Somewhere, probably not far beyond 400 kHz¹⁴, lies a sample rate which is capable of capturing all the nuances

of music that we treasure and hope to preserve. These are going to be some pretty big files: if we posit a 24-bit sample depth and a 192 kHz sample rate, we're looking at 4.15 gigabytes of data for a 60-minute program. The physical distribution media for this level of data density already exists, the DVD. Of course, MP3's and streaming audio formats will face problems an order of magnitude greater, when trying to figure out how to handle these files.

The bad news is that the next stage in degradation of quality for convenience' sake is already closing in. The Personal Digital Assistant (PDA) and the cellular phone are swiftly converging into a single device, with wireless Internet access trumpeted as the "killer app". Although the format of the Internet-accessed data will largely not change in the short run (streaming multimedia), the quality of the hardware pieces used to play it (post-Digital-to-Analog conversion), will, necessarily, be of very low quality. Imagine listening to the New York Philharmonic over the telephone. What is the point of this?

The stages of technological development in audio production and distribution have been presented here in chronological order, so that the shape of the trend will be apparent: a descending curve, which plummets steeply at the end representing present-day. Why are we here, at the low end of the scale? Music is not an instrumental art; totally abstract and containing no inherent references to anything in the real world, it exists only for the pleasure we take in hearing it, and for no other reason. Why deliberately remove so much of what makes music valuable to us?

I believe the current trends are anomalies, a temporary reversal of a broader historical trend towards *more* faithful reproductions, not less faithful ones. The reversal is the result of uncertain steps taken in the monumental shift from analog to digital forms,

and the even more uncertain steps in trying to figure out what can be done with the Internet. In the long run, continuing technological improvements in processing, storage and bandwidth will make higher quality aesthetic objects available in the digital realm. The markets will enforce, as they are already doing, the limits of what we are willing to accept for the sake of convenience.

COMPUTER ART

Reproductions of any kind, even the very best, will always be pale imitations of the real thing. Despite the convenient access to reproductions made possible by modern technology, art lovers are flocking to museums in record numbers to see the actual paintings. In fact, the increasingly broad access to reproductions of fine-art images may be creating a future audience, as users' appetites for art are whetted by the reduced versions they find on their computers. People are still willing to pay \$200+ for a pair of tickets to the ballet, symphony, and opera, so first-hand experience still counts. This area, first-hand experience, is where technology will eventually have its most interesting and transforming effects.

Although computers are used now almost exclusively as platforms for displaying art objects whose original venue is elsewhere—the art gallery, the concert hall, the film theater—computers have a largely untapped potential to create aesthetic experiences which are brand new, experiences which are inherently tied to the computer platform itself, and which are logical consequences of the unique structures and capabilities of the computer. Instead of acting as a secondary venue for diminished forms of traditional artworks, the

computer will become a primary venue for artworks which are native to the machine. Not only sounds and images (which have been used in combination in film for decades), but also multiple logical paths, and most of all, user interactivity, will be the materials from which not just new artworks but new forms of art will emerge. Already one possibility is beginning to be dimly glimpsed, a quality known in the industry as “immersion”, which refers to the ability of a multimedia presentation to draw the user into a palpable sense of an alternate reality. Although at this point the term is mostly used to describe detailed PC games, when more serious subject matter is treated, a new artform will be born. Another, related form is in the slow process of continuing development. Multimedia artists who perform on stage, like Laurie Anderson, depend upon computers for both the production and display of the elements used in a performance. The content of such a performance can be anything at all, but the natural tendency of such pieces is to give a representation of the mind of the artist: the sights they’ve seen, in both moving and still images; the sounds they have heard, and create; the thoughts that occur to them, in the form of text and spoken word; as these elements accumulate in the form of a performance or as a piece of installation art, the effect is very powerful. Creating an installation instead of performing onstage allows the addition of interactivity, as the patron steps into the space of the work and finds it reacting to him. All of the hardware and software requirements for the creation of these new forms of art have already been met. Even automated production tools are already available to assist the artist in the creation and assembly of multimedia objects, so that not all the artist’s time is taken up by attention to purely technical issues. Interactive artificial worlds will eventually include touch as well, when the technology for force-feedback gloves matures.

At this point in history, technology has done everything to increase the convenience and ease of access of users to music. It has done almost nothing to assure the quality of the actual reproduction itself. In the mad rush to create Internet music applications, so much has been sacrificed that consumers are beginning to show resistance to any marketing of the latest, very low quality formats. They will accept such degraded performances when they are free, as they did before with cheap transistor radios, but will not pay for them. As Moore's law continues on its relentless way, issues of quality will eventually be overcome, but technology-based reproductions will never begin to compete with the originals in their aesthetic impact. The real potential of technology for aesthetic experience lies instead in its ability to foster art forms which are altogether new, whose intrinsic nature springs from the new possibilities that technology raises.

ENDNOTES

1. Styron, William. *Darkness Visible*, Vintage Books, January 1992.
2. SONY online -
http://www.supremevideo.com/audio/sony/mini_systems/mhcnx3specs.htm
3. Revox online - <http://www.reeltoreel.de/worldwide/A77.htm>
4. de Paravicini, Tim. Audio (magazine) January 1995, interview with Bruce Bartlett.
<http://www.ear-usa.com/timdeparavicini.htm>
5. SONY online
6. Weaver, Greg. "The Vinyl Word", Soundstage.com, 1998.
<http://www.soundstage.com/vinyl/vinyl041998.htm>
7. Electronic Musical Instruments
<http://www-ccrma.stanford.edu/CCRMA/Courses/154:1996/week3/wednesday.html>
8. Heend, Dan. "The 24-bit Field Recording FAQ" High Fidelity Microsystems, November 3, 2001.
http://www.24bitfaq.org/#Q0_3_5
9. Morgan, Michael, with Wandling, Jeff and Casselberry, Rich. "Webmaster Expert Solutions", 1996, Que Corporation
<http://docs.rinet.ru:8083/WebLomaster/ch34.htm>
10. Ibid.
11. Weaver, Greg, 1998.
12. Fraunhofer Institute – Audio and Multimedia
<http://www.iis.fhg.de/amm>
13. Bouvigne, Gabriel. "Listening Tests of MP3 Using Different Bitrates" MP3 Tech.org. 1998.
<http://www.mp3-tech.org/index.html>
14. de Paravicini, Tim, 1995.