Recording The Future

A call for a sensible, open-ended advanced audio standard for DVD

by George Massenburg

So what’s DVD, and what does a standard mean to me anyway?

Hardly anybody thinks about the creative community anymore when a new technology is being launched. With few exceptions, technological decisions seem now to be made in boardrooms (a.k.a. “artist-free zones”) and artists, or “content providers” as we are now known, are expected to fit creatively within the constraints of formats that are introduced for strictly commercial reasons.

This marginalization of artists has not always been the case in the evolution of audio technology. One recalls the important, but suspiciously oft-cited, example of Les Paul and his promotion of eight-track master recorders, and the audiophile contributions to classical recordings by some great conductors – Leopold Stokowski, Herbert von Karajan and Erich Leinsdorf come quickly to mind. Not to mention the importance that producers like Walt Disney placed on improving sound reproduction for film for purely artistic reasons. But recent trends have headed elsewhere. And to many, the darkest moment in audio technology was the introduction of the digital Compact Disc, mostly for reasons of compromised sound.

Audio may, however, be about to embark in a new direction, built upon a new technology called DVD. This so-called Digital Versatile Disc is a higher-density permutation of the CD, and it promises to be the first true digital multimedia format, able to satisfy the demands of the video entertainment producers to replace the VHS videocassette, the computer and video-game community’s need for a much-higher-capacity CD-ROM, and the (by comparison much smaller) audio community’s interest in tagging along for a number of reasons.

The specification of an audio standard for DVD is a subject of great frustration for audio professionals, market-battered men and women who are spending a great deal of time and megabytes of e-mail in discussions that seem to meander from the arcane to the pointless. I think it’s time that those of us who make our living on the creative side of audio come to some sort of an agreement amongst ourselves as to what this new format could offer us all, and then concentrate on getting the attention of standards-writers. Will we be able to agree on anything? Maybe – maybe not. But if we can agree what not to agree on, I submit that a flexible specification is the obvious next step.

For those who may have missed the news, DVD is being pitched, and is currently primarily specified, as the next great video format. It looks a lot like a CD or CD-ROM, but it has far greater capacity - almost 5 gigabytes of data compared to the 700-odd megabytes of the CD - and a much higher data rate, and can hold a full-length data-compressed, digitized movie on a single side. Hollywood is hoping that DVD will bring more customers back to the video store to buy rather than rent their movies and also that this new format will deal a comeuppance to the now-powerful cable companies. Japan is hoping that DVD will lead more buyers back to their consumer electronics stores. And, in the meantime, Silicon Valley already has start-ups designing and manufacturing silicon hardware, and multi-gigabyte applications on DVD-ROM’s are close at hand. Where does that seemingly dwindling cadre of audio pros fit into this?

Simply put, we have an opportunity to utilize the DVD as a medium for a new generation of recorded sound.

But, first...

What are the major issues within an Advanced Audio Standard?

• Digital Sampling Rate
• Digital Word Resolution (number of linear bits)
• Digital Data Compression (or rather, bit-rate reduction; from “lossless” coding schemes to those which selectively corrupt data, such as “perceptual encoding”, et al.)
• Number of channels.
• Compatibility (especially backwards compatibility with the current Compact Disc)
• Intellectual Property Protection (Copyright issues)

Digital audio, again?

First, it should be pointed out that our universe is probably more “linear” and “continuous” than not. There are few examples of “digital” like things in nature, with some possible exceptions: the protein “bits” in DNA that make up the genetic code, and quantum effects at a sub-atomic scale. It is more a world full of ambiguities and chaos, but rich in fine detail and continuity. It is not an “analog” world, either. May I submit that nothing in nature closely resembles our current digital coding of sound.

But, ignoring the sentiment of some audio professionals that digital should just go away (and, mind you, sales of 33!/2 most-often-black polyvinyl-chloride record albums are up in the past two years, especially in Japan), many see some of the benefits of digital media, and very few can ignore that it is already virtually impossible to circumvent digital conversion at some point in getting sound to the public (except for the LP). We are well-advised to seek out better digital. Unfortunately, there is currently no consensus as to exactly which digital format is optimum - there is even disagreement amongst us over how to weigh the factors within the constituent issues. There are even a few among us expressing satisfaction with the current 44.1kHz, 16 bit format of the Digital Compact Disk (or “CD”).

Some 16 years ago, at its introduction, the CD was touted in the pages of Billboard as “…[embodifying] perfect sound, forever!” Did we ever really believe that? Many of us listening to the first CDs on early players had, at best, mixed feelings. The sound of many of these first CD’s was cold, “flat,” hard-edged and thoroughly foreign. The first digital recordings that most of us made seemed to suggest that everything about working with digital was in many ways different than working with analog – that few, if any, of the analog recording tools upon which we had come to depend were anything but obsolete. Not only that, the methods that we had to measure the new gear weren’t explaining problems that we could identify. Well, folks, it wasn’t perfect and it didn’t last forever. It’s ready to be eclipsed, technologically. And if there’s a lesson here, it is, “Standards in our time tend to have unambiguously rock-solid, brick-wall limits.”

Oh, and besides, none of it may matter to music, anyway...

As we approach possibilities of DVD, I think we have to separate the audiophile business from the rest of the music business. Producers and some artists would like to believe that all consumers and all A&R executives appreciate better sound. But as producers and recording engineers, it is important for us to remember that while we perceive our job as pushing the envelope of the business, we must realize that it is most assuredly not the main business.

If fact, most of us have learned the hard way that “high fidelity” has very little to do with selling records. In his keen and perceptive article in the November 24, 1996 issue of the New York Times, Lawrence B. Johnson states, “Some would insist that it is not so much the sonicly veiled voice as the music itself that stirs the community’s interest in tagging along for a number of reasons.

The specifi
from vinyl to CD? With DVD, we will move yet again and I believe we will move to a more ambient recording environment. This could have a profound impact on popular music as we now know it.

Practical issues abound.

What if a DVD audio standard required a double-sided disc? That would certainly get the attention of the record manufacturers – they’d have to rethink where to put the darn label. And what about compatibility with existing CD players? Producers and artists may be excited by DVD’s possibilities but at the same time it may leave the record labels cold.

DVD audio technology is being challenged by the realities of selling records, the limitations of today’s production tools, and our creative inexperience, to prove its relevance. Imagine if we as an industry debuted the future of audio technology and everybody yawned?

So we may have to take reality into account.

Sample Rate and Word Width.

About ten years ago, in the middle of the digital dark ages, I had the pleasure to spend some time with Dr. Roger Lagadic who, at that time, was director of Sony’s AP Division with responsibility for professional digital audio products. As he began to drive me from Shinjuku (in downtown Tokyo) to Atsugi (in Kanagawa about an hour and a half’s drive away) I mentioned that it would be nice to see some audio products use a significantly higher sampling rate. For the remainder of the trip I endured lecture, discourse, and a lot of griping about how irresponsible it was to so lightly conclude that such a fundamental change in technology was a priority. Sony’s, and almost everyone else’s, world, was (and, in large part, still is) built on 44.1kHz, and the impact of and ramifications to making this sort of change is intimidating.

There are no issues more likely to spur discussion to violence than where to draw the line for digital word width and even more so for sample rates. The costs involved in retooling for the latter are, to say the least, significant. Worse, there is little hard scientific data (and seemingly only anecdotal empirical data) in psychoacoustics to support the usefulness of signals beyond 18kHz and resolutions in the 20+ bit range.

Even so, as I mentioned before, many of us welcome the opportunity to go beyond the 44.1kHz / 16 bit world of CD’s. Many of us have, by now, heard the benefits of increasing resolution (most often 18 to 20 bit A/D and D/A conversion at conventional rates) as well as listening to increased sample rate (Allen Sides and I owned and used Mitsu X86HS 96kHz /16 bit machines). We like what we hear - most often we hear improvements in a sense of space and depth, or more ‘air.’

As an engineer who has lived in front of the monitors for a great many years, I wish to say that I often spend a lot of time with a full-range sound in the sense of a wide dynamic range, as well as a wide frequency-response (when I’m working analog). For the weeks, months, or years that I have lived with a given project. And I can only add that I am more often disappointed than not with what comes back from CD replicating facilities. Worse, it’s dawning on me that even as the possibility of a much better digital product format comes closer, the format of my digital archives (fixed like almost everyone else’s at 44.1 to 48kHz, 16 to 19 bit) severely limits their usefulness. Why, oh why, haven’t we digital-format producers been archiving our masters at some much higher standard than our release media?

Assuming that I do want to put out more life-like recordings (and I’ve had a pretty fair response to the recordings that I’ve made over the years), let me tell you that I’ve never heard the recorded sound of a solo violin or a trumpet with a Harmon mute come to close to an original performance. And, look, I’ve heard the arguments about single-variable tests and questions about elements of the signal chain other than the converters. I don’t buy it. During the often-strident arguments about the contribution (or lack of same) of so-called out-of-band high frequencies, I feel like I’m hearing a lot of, ‘Pay no attention to the man behind the curtain.’

Warning! Personal opinions follow. Flame off!

I mixed a record several months ago that came off analog multi-track tape; it was the first time I’ve worked from an analog tape master to an analog tape mix in a long time. The artist and producer preferred to record that way and I was working just as an engineer, but I learned a lot from the experience. It was really nice hearing a real hi-hat again after 15 or 16 years. It actually sounds pretty good. C’mon guys. We know that digital is predisposed to harshness - it is not the same listening experience as we came to know with analog (ignoring for the sake of argument there are certain attractive quantum improvements in noise elimination and media integrity). Digital harshness and stickiness are okay if what you want is a provocative and irritating (stimulating?) sound, but that is not the record that I’m going to put on for a spiritual experience; I’m more likely to put on one of Doug Sax’s acoustic music CD’s, or Gabe Wiener’s early music CD’s, or a Peter Gabriel these days. I want something that can I can feel safe with. Doug’s and other’s fine work and great efforts notwithstanding, I love the lace and ligne of a virtuoso performance and I do not believe that 16/44.1kHz is best in delivering that.

Like pioneers of yore in our industry, I would prefer to experiment on my own thank you very much. And the first thing that I’m going to do when I have cobbled together enough advanced technology is to try it for myself and hear how it feels.

And that technology is becoming more accessible to us. We have a box from Prism that allows us to record 6 channels of 20 bit data (albeit at 44.1kHz) on a DA-88. Sonic Solutions is said to be shipping 96kHz software for their workstations (and they are committed to wide words). And, the newest Nagra D records 96kHz with 21 bit resolution.

End of opinion. Flame on.

There are several sample rate / resolution options before us. One could, in fact, hold 44.1kHz sampling and improve the resolution to 20 or 24 bits. One could hold the resolution at 16 bits and increase sampling rate (although increasing word width comes “cheaper” than increasing sample rate). There is interest in an intermediary 64kHz sample rate.

Gabe Wiener puts the issue thusly: “The ultrasonics issue as I see it is this: a) We know there’s stuff above 20K. b) We know that the 96 kHz converters sound better than the 48 kHz converters. c) We don’t know if they sound better because 1) having a brickwall filter, oversampled or otherwise, at 22 kHz is just not a healthy thing to have, or else 2) the filtering iso facto isn’t doing the damage, but the missing ultrasonics are. The big mystery is -- Does 96K sound better because the wider bandwidth makes filtering easier? Or because it produces more harmonics? Or both?”

In any case, we seem to be slouching towards a 96kHz (88.2kHz) 24 bit. The ARA and others back this direction.

A completely different format, the high-bit-rate single-bit stream (two so-called bitstream coding methods from both Sony and Philips) are departures from so-called PCM. One real advantage of the Sony direct-stream digital, or DSD, approach is that it is rather straightforward, and often trivial, not only in converting to analog, but also in converting to other common digital data rates from the original single-bit stream. With that in mind, there are some problems, notably the significant reduction in data storage efficiency.

In any case, there are real questions regarding what is audible and what is an excessive, wasteful use of bandwidth. Clearly, we need more substantive testing and evaluation, but, again, flexible standards for the medium would do much to facilitate comparative analysis.

Issue: Compression, or bit-rate reduction.

Data compression is a wild card.

All current data compression schemes that cannot be classified as “lossless” modify, or corrupt, program material in some way. Lossless compression methods in digital audio are identical to data compression techniques in computers - you retrieve data identical to before the “compression” process. I don’t like “lossy” compression and I personally don’t know of an audio pro who dares to admit that they prefer the sound of data-compressed digital music. But bandwidth and capacity do not come cheaply, and data compression is certainly something we have to live with.
Currently, the data compression scheme that has been designated for the audio channel on video DVD’s is Dolby AC-3. Dolby has tremendous clout and they always seem to be there when standards are being decided. However, I would hate to see Dolby walk away with more standards, because AC-3 is insupportable as a high-fidelity standard. Dolby has walked away with some real trophies, such as an audio standard for HDTV. I grant Dolby the right to be one of a number of encoding formats for DVD but I will oppose their right to be the “only” format, especially as time and opportunity brings more and better innovations to the field.

There are several alternatives that are clamoring for consideration. The Acoustic Renaissance for Audio proposal for 24-bit/96 kHz/5-channel audio, for instance, limits itself to only lossless compression. There are certainly degrees of audibility in “lossy” compression: DTS, Philips, JVC and Pioneer all present alternatives, they sound different, and producers should be able to choose the most appropriate method for their purpose.

Unfortunately, the DVD standards that are being discussed do not allow for the variable-bit-rate encoding of the audio stream (perceptual coding methods invariably utilize a constant-rate stream), thereby making lossless coding (with it’s unpredictable bit rate) very difficult to utilize efficiently. One would therefore hope that variable-bit-rate encoding of the audio stream find it’s way into a specification.

**Multichannel audio in the new millennium.**

It has been proposed that the greatest single commercial aspect of an audio-only DVD is multichannel audio. Here are some reasons:

- Multichannel sound is right in line with the evolving marriage of video and audio. Already, there are reportedly 25 million ProLogic-licensed units in place. There are in excess of 100,000 Home THX units as well. It’s easy to imagine there are a ready (perhaps even hungry) market for multichannel music.
- For strictly commercial reasons, multichannel sound will be appreciated to a proportionately higher degree by the ‘average’ consumer. I suspect that the vast majority of the population will have trouble identifying the difference between 16-bit versus 18 or 20 or 24-bit conversion. But boy, doesn’t everybody hear the difference between mono and stereo? And one can help but be impressed with the differences between stereo and 5-channel.
- Multichannel audio gives us a completely new “color” to put on our palette. It is by no means yet clear what stories are there to tell in a three-dimensional audio space, but there are certainly new possibilities.

What about the disaster of quad, you may be asking? Is the music business about to replay the biggest commercial flop in the history of audio, only now in DVD digital? I think not. Quad failed for basically two reasons: (1) there was no sturdy multi-channel medium (very little hardware and not only formal incompatibilities but any number of really bad solutions such as matrixing systems) (2) the software never matured - evolved practices never had a chance to happen. An uninformed industry had leap forward only to gradually learn that there were major drawbacks of matriced quad. The industry became nervous and further investment was withdrawn.

With DVD the dream is to establish a single, compatible format. And, during the course of the twenty years since quad crashed and burned, the technologies and techniques for creating multichannel audio have evolved considerably. The tools for multichannel recording are certainly available. One real limitation is that up to now there has been little demand for multichannel audio recordings. And, as a result, outside the film and video suites, very few music studios are now equipped for multichannel recording. It’s simply not trivial to record and mix 5 channel music in a two channel room.

Fortunately, once the demand arrives the move toward multichannel music studios will move rather quickly, since the tools already exist. Multichannel storage will be solved by the same storage devices that have fostered the project studio: the modular eight-track multitracks, as well as non-linear hard-disk and optical-based systems. The additional technology we’ll need is a way to move signals around. You could do this within the current construct of regular consoles but it’s pretty hard to pan three channels with a two-channel panner. Until music rooms begin getting requests to do music for video they don’t need to address this issue of multichannel panning.

There have already been several noble experiments. Pink Floyd’s tours, “The Wall” and subsequent tours, for example, explored the ability to tell meaningful musical tales in the front and back as well as side to side. James Guthrie has struggled with Q-Sound to capture some of these spatial ideas in two channel recordings. Although many of the tools aren’t quite there for the journeyman producer to really explore the medium, they will undoubtedly be coming down the line shortly. Certainly, film mixers have been able to mix for 5 discrete channels for some time (and sometimes, 7, with 5 in front and two in the back, or 5.1 feeding something into an omni bass boost channel). There are consoles that are designed to pan for the new format, so it’s all going to be available soon. I don’t think the tools, or lack thereof, is the biggest obstacle to our move toward multichannel music. I return again to the creative question — more information is wonderful, but what does the artist do with it? And the answer is probably, “Give them a tool and find out.”

**Lick is at least three-dimensional, and, in the sense that we use music to tell stories about life and sound to paint sketches, multichannel technology will give artists a new palette for the telling of stories in multiple dimensions instead of just left and right. Other than that we don’t know very much. We do know multichannel music is going to be different than sound effects in music and sound effects for film. Early on we realized that you couldn’t utilize the full 360-degree sound space for music because there was an ‘exit sign’ phenomena. If you had a lot of audio going on in the back, people would have a natural inclination to turn around and stare at the back of the room. However, there are some good musical indicators that demand closer examination.

**Remix, Then and Now.**

Linda Ronstadt’s “Cry Like a Rainstorm,” which we did in 1989, is one example of a record I wish had been recorded for multichannel release. At the time, we recorded the album with a lot of ambient information and, when DVD comes along, I can’t wait to remix it for multichannel and put some of those sounds back in a larger context. A lot of the sources were recorded tight with two separate ambiances. It’s not impossible to imagine something originating in the back and then folding up to the front. I don’t know how it will actually unfold, but I can’t wait to find out. It would be very nice to explore that extra space. When we make the move to DVD you’ll see a lot of remixing going on. And, as DVD gets closer, you’ll see artists beginning to approach their current projects with a future multichannel release in mind.

Take one of my favorite artists for coming up with new ideas — Beck. I love what he does as he explores the boundaries of technology in a way I frankly haven’t heard since the Beatles. At this time trouble with $100,000, distortion, saturation, speed and (shhhhh) abusing equipment! I can’t help but think that somebody like Beck, given more speakers to work with, will find something very interesting to say with multichannel sound.

This is the most critical issue. Is this not more an artistic challenge? If, in fact, there is a multichannel future for this format and we are given the right tools, then it is up to someone, somewhere, working with this bigger palette to come up with a work of genius to show us some ways creatively. One really wonders whether the vast DVD market really needs a higher-fidelity specication. Will incrementally better sound quality yield a better musical product?

And while we’re on the subject of number of channels, there are some other options to be considered. Tomlinson Holman, citing the work of Michael Gerzon and his own tastes, would prefer to have the ability to store more channels (as opposed to better frequency response or wider digital words), perhaps as many as ten. And there are other options as well, many of which are possible to imagine in a rich and evolved, and, importantly, long-term future.

**So?**

What is difficult to imagine right now is a flexible system for playing all of these formats back. How, for instance, is a consumer with a two-channel 44.1 kHz playback unit going to be able to play back a DVD? That may be solved by the record companies wanting to keep the inventory down to single unit or single format inventory and asking for the so-called “Red Book” layer for standard CD playback on the DVD standard. It would accommodate an arrangement where you could either play this new DVD on a standard CD player or you’ll put it into a DVD player and the laser would focus on the layer below Red Book by simply looking through the Red Book layer. Sounds logical, but how would the player be told whether it would be a two-channel or a 10-channel surround sound mix? That’s why, regardless of what format they decide on, each and every DVD will need to contain a format block or a parameter block, a descriptor that would tell the player everything about the disc you are about to play. It would need to provide enough room to describe all the possible encoders and decoder formats — and that’s pretty tough.
At the same time, the electronics manufacturers are going to be looking for ways to keep the cost of the equipment down. Well, if you’re reducing the cost of the equipment, you’re unlikely to find playback automation on board which would require some signal processing (which is always getting cheaper but is by no means “free”).

**Is this curtains for Mr. CD?**

You see, we can get caught up in all the technology without considering the realities of how the consumer is going to access all this new audio data. Will he/she be using a dedicated player, a home theater, a home PC, some sort of hybrid cable/internet box, direct broadcast satellite, or what? I think, for the time being, we are going to have to consider one box and this box will likely be a video playback unit. (The days of a dedicated CD player, at least outside of the audiophile community, is coming to an end.) Once Hollywood solves its issues about copyright protection, which held up a video DVD introduction for some time last year, the first box is likely to be: a video DVD playback unit that can playback constant bit rate video JPEG and AC3-audio with 5 channel plus low frequency support. The audio community will have to ride on the back of the beast. That’s not the optimum creative decision mind you, but it is likely to be the most realistic one. What does it mean? It means that initially artists and producers will need to create music on a video DVD recorder without video, with the ability to move it directly to 5.1 without AC-3.

What does this mean for the studio community? It means that many of our new techniques for DVD production will come from the film soundstage. I’m already seeing music and film mixing coming closer and closer together. For example, I recently did a Tom Hanks movie. I did my CD mix and then we broke it out to Tascam DA-88 and we looked very carefully at what we needed in the dubbing theater. I needed to break out the reverb elements so that if he brought all the faders up to zero he’d get my mix, but he would more than likely pull back the reverb components so he could get a little more clarity for the film mix.

In other words, like it or not, we are already slouching toward a DVD standard. What we would like to see beyond that is something that will accommodate future technologies. What if, for instance we discover that the great majority of the public doesn’t buy into the 5-channel 24-bit/96kHz format? Wouldn’t it be great to be able to do something else with this new platform? Let’s not get bogged down because DVD is primarily video-biased and must deal with multi-continent video standards and compression schemes that are far more difficult to deal with than audio considerations. DVD audio allows for multiple formats to coexist and having enough imagination right now is important. Let’s simply imagine where we can go in the next 20 years. Is there anything we’ve imagined that we haven’t covered in this new format?

The most important thing to remember is that today’s commercial record business is not driven by audio. It was for a short time during the Seventies but no longer. Remastering the back catalog for CD was a bonanza for the record business but we’re talking about something quite different in terms of remixing for DVD. Sure, artists like Pink Floyd are going to do it. But we’re talking about a tremendous new investment required here (not that I would rush out and open a new mix room in L.A. just yet.) Remixing just one major release for multichannel could mean 6, 8, 10 months or a year of work, so the record company is going to have to have a promising return on investment to warrant putting that kind of money in a catalog property. And, in the meantime, you’re not going to find many major new artists exploring the possibilities that these new standards will provide. That’s simply not the nature of the music business today. Today’s biggest artists are not, shall we call them, “hi-fi artists.” What would alternative music producers, who are for the most part classic-reductionists, do in multi-channel?

**A modest proposal**

Do we really believe that the DVD or any quantum leap of technology will be a panacea?

I’d like to think so, but I’ve got my doubts, mostly about an inflexible standard. Then again, for us to work together as an industry, we have to have agreements, and if I was the one sitting up in the boardroom deciding how to launch this new audio format — balancing my commercial responsibilities with my creative passions — I would probably take things one step at a time as is being done right now with an initial standard for movies. We have to work from a starting point and then expand the format as technology matures, costs come down, and exotic technology becomes commonplace in the future. My opinion is that 96kHz, 24bit, “lossless” coded 5.1 channels is a good next step. But that having been said, we need to establish the most flexible format set possible. In fact, I’d like to see the ability to have an open-ended capability so that a producer can say, “I want two hours of sound and I want it in 5 channels. I don’t care so much about the sound quality, so let’s compress the hell out of it or, maybe I’ll change my mind half way through the project so that I can decide to improve the sound to just two channels of extraordinary 200 kHz bandwidths.” The producer and artist should have a format that can be as flexible as their imagination demands, and as adaptable as future technology will be in the foreseeable future.

**Acknowledgments:**

The author would like to acknowledge the help and contributions of Peter Asher, David Goggin, Tomlinson Holman, Kyle Lehning, Richard Mithun, Anthony Morris, Richard Pierce, Doug Sax, Chris Stone, and Gabe Wiener, all of whom added in different, substantial, and often profoundly important ways to this piece.