Institute of Acoustics

A SURVEY OF THE OPERATIONAL CHARACTERISTICS OF TWO-CHANNEL DIGITAL AUDIO RECORDING SYSTEMS

F.J. Rumsey

University of Surrey

INTRODUCTION

The use of digital audio recording equipment is now widespread, yet a number of formats exists and standardisation is slow in coming. It would be unrepresentative of the state-of-the-art merely to compare the relative merits of formats without being product-specific, as formats only dictate factors such as sampling rate, tape speed and tape format, having no control over operational factors such as interfacing between machines, metering, and editing equipment. It is these more practical and operational considerations which this paper sets out to discuss, with reference to a number of commonly encountered two-track or stereo recording devices.

Currently, usage of two channel digital recorders falls into two fairly distinct camps: the so-called domestic or semi-professional devices, so far represented almost entirely by Sony EIAJ-format equipment such as the PCM-F1, and the fully professional systems such as Sony's PCM-1610 or 1630 and Mitsubishi's X-80 or recently X-86 reel-to-reel machines. As there is at least an order of magnitude's difference in cost between the two types of system, it would be inappropriate to make certain direct comparisons between facilities found in each, yet there are a number of areas of interest which warrant discussion.

PHYSICAL FORMATS

People wishing to make digital recordings have the option of doing so on a number of physical media which differ widely in size and convenience. It may be noted that an hour of two-channel audio cay exist on a large and bulky U-matic cassette, a tiny, light R-DAT tape, and a ten-inch open reel. Each of these may be more or less susceptible to damage, and more or less convenient in use, but each may store two channels of digital audio at the same sampling rate, with the same number of bits per sample, and thus with a similar potential audio quality.

Open reels

Recording engineers are used to working with open reels. This alone could be one major argument in their favour, though the most important one is the ease with which splice-editing may be performed.

Two major open-reel recording formats are currently in evidence, these being the so-called PRODIGI format, and the DASH (digital audio stationary head) format, although the late appearance of real machines like Sony's PCM-3402 (DASH), Studer's D-820X (DASH) and Mitsubishi's X-86 (Prodigi) has meant that the industry so far has little experience of them. Prior to this, a stereo machine from Mitsubishi (the X-80) and its Telefunken equivalent (the MX-80) had been used in limited numbers, although these machines did not conform to the Prodigi format and tapes are not compatible with the Prodigi format.

Open reel recording of digital audio has very little in its favour, except that it allows splice editing of the tape, which for reasons of economy and tradition is a feature that has been insisted upon by recording studios and broadcasters alike. Consequently, formats were devised and machines designed to ensure that some provision was made for the cutting and splicing of a tape which can ill stand such physical handling due to its delicacy.

Because of the tape's thinness, a standard NAB reel allows around one hour of recording at a speed of fifteen inches per second, and around two hours at seven-and-a-half inches per second. Sony's PCM-3402 DASH machine will run at either speed (the higher speed allowing ease of editing, and the slower speed giving the benefits of long storage time), while the other machines allow only the higher speed (with the exception that the DASH format dicates a slight change in tape speed when using the 44.1kHz sampling rate instead of the 48kHz rate). This makes the storage time currently available on open reels comparable to that which can be achieved with cassette formats, especially if larger reels are used. It also compares well with the storage time available with analogue machines.

From a negative point-of-view, tape on open reels is clearly vulnerable to more damage than tape mounted in cassettes. Experience so far indicates that the thin digital tape is more prone to edge damage, and that edge-damage is exacerbated by the "leafing" which tends to occur when winding thin open-reel tape at high speeds. For this reason, certain manufacturers recommend high-precision spools and slow winding speeds. Both formats ensure that auxiliary and cue information is recorded on outer tracks, so that these will take the brunt of any damage.

The narrow head gap and short recorded wavelengths involved with stationary head recording give rise to severe signal loss with even extremely small head-to-tape spacing errors, such as might be caused by debris or dirt stuck to fingerprints on the tape. This can occur only too easily when everything is exposed to the air, when tape is handled regularly, and when splice-edits are performed, resulting in erroneous replay which at worst can cause complete dropout and at best may be disguised or corrected by the error correction system. Considerable efforts have been made by the designers of both DASH and Prodigi formats to ensure that the error correction employed is capable of protecting the recording against these horrors, and now that the so-called "double-DASH" format is being used by both Sony and Studer at 15 ips (which effectively doubles the recorded redundancy) recordings made on all the machines mentioned above appear so far to stand up to a respectable amount of abuse. The only exception to this is in the splice-editing situation which will be covered later.

Cassettes

Many of the aformentioned dangers of physical damage are obviated when the tape is in a cassette, but immediately one loses the facility for cutting the tape, requiring that some form of electronic editing is used. A number of cassette formats are commonly used, the majority tied to video standards such as Betamax, VHS and U-matic, while the recent DAT format uses a dedicated tape.

So fár, all of these have used a rotary-head recording method, which has allowed slow linear tape speeds, and thus a relatively long playing time per inch of tape. Coupled with this, the EIAJ format used in Sony's PCM-F1-style processors was designed for use with domestic video tape recorders which results in low overall costs, both of equipment and of tape: it is possible to gain three hours of digital recording, having the same potential audio quality as a professional recording, for only five pounds or so, and it is doubtless that this has contributed largely to the format's success in professional circles.

The cassette formats which are video-based generally require that two pieces of equipment are needed for recording: a processor to convert audio into numbers and thence into a video format, plus a video transport to record the result, whereas a dedicated format like DAT will usually reside in a single unit which incorporates the transport.

Cassettes do not always imply a long recording time: indeed the U-matic tapes used with the professional CD mastering system last a maximum of seventy-five minutes, just long enough to fill a compact disc, yet they are exceedingly large compared with a DAT tape. Partially this is due to the technology available at the time when the Sony PCM-1610 system was designed, and the need for a robust format with lots of redundant data. Even so, engineering has advanced far enough during the last five or six years to provide apparently better error correction and an equally robust format with comparable audio quality on a tape with a fraction of the size.

RECORDING AND PLAYBACK

Stereo and two-channel formats

It is important to distinguish between a true two-channel format and a stereo format, when they sound as if they mean the same thing.

All the cassette formats so far available are stereo by nature: in other words, they allow recording of both channels at the same time, but not separately. Both stationary head formats allow separate recording of the two channels, so that each channel could contain a completely different monophonic recording, and could allow independent punch-in and out. This facility may only be important in a limited number of applications, but it is worth remembering. Also, it should be noted that the original Mitsubishi X-80 was a stereo machine, not a two channel machine.

Punch-in and out

It is only on the open-reel formats that the facility for "punching-in" to a recording is quoted as a feature, yet the electronic editing process used for the cassette-based PCM-1610/1630 system relies on the ability of the edit controller to perform a similar function. The difference lies in the ability of the open-reel machines to perform a cross-fade at the punch-in point, with another at the punch-out point, so that a seamless join with the old material is apparent, whereas the most that a video recorder can usually do is drop into record mode at the frame boundary. A drop-in using the latter method might produce an audible glitch if the signal waveform were to jump considerably at the join, even if the digital data stream was not corrupted. Most of the open reel machines allow for internal variation of the crossfade time at the in and out points.

Some U-matic recorders allow for a pre-programmed drop-in, which results in a frame-accurate edit between old and new material. Whether or not this produces problems depends on the digital recording format in use, because the EIAJ format, for example, cannot be treated in this way due to the fact that the interleave pattern of data is not complete within a video frame, thus giving rise to errors if one frame is joined to another completely unrelated frame. The PCM-1610 format, on the other hand, does not suffer from this problem, allowing unrelated frames to be joined without digital errors, although any crossfade from old to new material must be performed by external processing, such as occurs in the associated editing system.

Punching into recordings is primarily a feature of multitrack work, being a relatively rare practice in stereo recording where it is of most importance in electronic editing of the mixed material. This will be discussed in the relevant section.

Robustness of recordings

The most important thing about any recording is its reliability. No matter what the audio quality, or how easy it is to edit, the recording is no use if it has dropouts or audible errors. A format's ability to cope with errors is often referred to as its "robustness", but this will often be severely taxed by factors such as poor tape, or mechanical alignment of the recorder. As an example, it would be interesting to look at the problems encountered with EIAJ-format recording, which although used widely in professional circles, has shown varying degrees of reliability over its life so far.

Although the EIAJ standard was first designed to handle fourteen bit recording only, Sony pushed it to the limit in the PCM-F1 by removing some of the error correction data to make room for another two bits of audio, so as to bring this "domestic" format up to sixteen bit "professional" standards. Consequently, almost everyone who uses the EIAJ format now uses it in its bastardised sixteen-bit form, which has not the error correction ability of the original format. Operational experience has shown that at best, sixteen-bit EIAJ recordings are reliable and may play for hours with no dropouts, yet at worst they may be plagued with dropouts, spits and pops, and interpolation errors. Interpolation errors will not always be obvious as such, unless the processor is fitted with error status indication lights such as are provided by Audio and Design in their "professionalised" PCM-701, thus many people will not be aware that such problems exist, as the only indication of error status on a standard processor is a PLAYBACK MUTE light. The audible results of persistent and prolonged interpolation errors will be a reduction in audio bandwidth.

A number of factors have been seen to affect the performance of EIAJ format systems, many of which may have been attributed falsely, and it is still not entirely clear what affects the error rate the most. Nevertheless, it seems clear that the performance of the sixteen bit version of the format is often very close to the edge of acceptability. The better video tape transports have been found to give significantly better performance than cheaper ones, as in some cases has the use of high quality Betamax tapes, and recent information indicates that the new "HQ" VHS machines improve the error rate enormously, although VHS tapes are non-standard in this application. Many people have resorted to the use of U-matic machines instead of domestic machines with EIAJ processors, and this increases the reliability at a price, as well as a loss of recording time.

Very few professional users rely on a single EIAJ format recording for important work, especially when using Betamax transports. They will be more likely to run two machines in parallel, or use U-matic recorders, as it is most unlikely that a dropout will occur in the same place on two tapes. It remains to be seen whether it is worth any further research into the improvement of the reliability of this format, especially now that DAT provides similar facilities in a far more robust format.

Level control and metering

Digital recording presents the user with a whole new set of problems in setting levels and metering the signal to be recorded. The electrical levels within a conventional recording studio and the metering used are often far from ideal for making a high quality digital recording, unless the engineer is aware of the pitfalls.

In order to make a high quality digital recording it is important for the peak recorded signal to make the maximum use of the dynamic range available: in other words, to use all the bits. Many people are not aware that digital recordings which only peak half way up the meters are only perhaps twelve bit recordings, even though they may be using a sixteen bit device. To drive a professional processor like the PCM-1630 to peak recording level requires more than +20dBu, which is extremely high when compared with most studio reference levels of around +4dBu, and requires that the needles on the meters of the mixing console are permanently on the endstops in order to produce enough drive. The original PCM-1610 further perpetrated the problem of low recording levels by putting a 0dB mark about 20dB below the peak on the meters, encouraging users to treat it like an analogue recorder and peak around the 0dB mark, thus producing a thirteen bit recording.

Mitsubishi's X-86 incorporates two sets of meters: a pair of mechanical VU's, and a pair of peak-reading bargraphs. The 0VU point occurs at +4dBu, as is conventional practice, but the 0dB point on the peak meters occurs at +18dBu which is the peak recording level, but which is bending the needles of the VU's. This brings into question the sensibility of using VU meters for digital audio, as they are only a satisfactory indicator of steady-state levels. A VU meter's reading of peak levels is abismal, often showing as much as ten to fifteen dB lower than a peak meter on certain programme material, which would result in regular overload of a digital recorder on peaks if the VU meter were trusted. For digital recording (and, in my opinion, for analogue recording) it is necessary to know exactly what is going on with audio levels, as it is vital that the recorded level is controlled to be as near to the peak as possible, yet without going over it.

The PCM-F1 and its counterparts operate at lower electrical levels, as befits a domestic device, requiring not much more than 0dBu to drive them to full recording level. They are also provided with reasonable peak-reading meters, having a sensible scale covering a range from around -50dB up to 0dB at the top. If these are to be driven from professional mixing consoles then often the levels have to be held lower than usual, unless the digital processor has had one of the professional interfaces fitted which allows the unit to operate at higher electrical levels and provides it with a balanced input and output on XLR connectors. The Audio and Design modifications to the PCM-701 allow a switched variation of input level in 2dB steps to accomodate a range of standards.

One upshot of this problem is that unless the mixing console is fitted with accurate peak reading meters which are capable of showing up to the maximum recording level on professional processors, the meters on the processor itself will have to be used more often, and these are often a long way from the desk. In order to alleviate this situation in one instance, a remote meter has been designed and built by Totalsystems (a private audio design company) which plugs into the status connector on the PCM-1610 or 1630, and which can be mounted at the mixing console.

Metering and reference levels require some new thought in the age of digital audio if the best use is to be made of the equipment.

Sampling rates

It is perhaps unfortunate that two professional sampling rates have been established, but this is yet another example of the inability of the audio industry to apply standards and adhere to them. Although the 48kHz rate has been stipulated as "the professional sampling rate", many people find this highly inconvenient because they are making CD masters which have to be at 44.1kHz, thus they work normally at this rate. Coupled with this is the fact that EIAJ processors all work at 44.1kHz, and that these recordings are often copied digitally to professional systems.

Consequently, both DASH and Prodigi machines offer the option for recording and replay at either sampling rate, usually changed by an internal switch. On the DASH machine this results in a lowering of tape speed for the 44.1kHz rate, in order that the density of data on the tape remains constant, and the setting of a flag in the recorded data which is used to switch the machine automatically to the correct speed on replay. The Prodigi format, on the other hand, does not require that the tape changes speed to accomodate the lower sampling rate, but still places a flag in the recorded data to indicate the rate used. An interesting side-effect of reducing the sampling rate without changing the tape speed is that the recorded data density decreases at the lower sampling rate, thus one would assume that the error rate must be lower also.

Mitsubishi's X-86, as the only Prodigi two track, does not switch the sampling rate automatically on replay, even though the machine clearly detects that the wrong sampling rate is in use and displays it in the form of a flashing warning light. The result is a shift in the replayed pitch of the music of a few semitones, which may only be noticeable if the operator is familiar with the music, or possesses perfect pitch. Switching back to the correct rate restores the pitch.

The R-DAT format allows for three sampling rates, although the domestic machines so far brought into the country record at 48kHz when using the analogue inputs, and disallow 44.1kHz to prevent digital copying of CD's. Professional machines do not prevent recording at 44.1kHz which brings an interesting new twist into the tale, as no longer will people be able to make CD masters on domestic DAT machines, whereas they are used to being able to make them on domestic EIAJ-format machines. Some domestic models are fitted with a Sony/Philips digital interface for digital transfer of recordings both in and out, but these will be of little use unless transferring to a professional device switched to 48kHz. The alternative is to use a sampling rate convertor, and it remains to be seen whether domestic sampling rate convertors become a feasible proposition.

Relative unimportance of the replay chain in digital recorders

A factor which few people consider when assessing digital recording equipment is how little the quality of the replay convertors matters when using the equipment to make compact disc masters. The latest trend in professional devices is to install domestic-style oversampling D/A convertors in the replay chain for higher sound quality, whereas these convertors never get to affect the sound of the compact disc as the digital recording on the tape is transferred numerically to the disc without passing any analogue conversion. Thus it is really only the A/D convertor quality which has any effect on the sound of the CD.

This makes so much nonsense out of direct comparisons between digital recording devices on grounds of sound quality, as the comparison will only tell the user what he will hear in his own studio, rather than what the listener at home will hear of his recording. In order to assess the sound quality of the important A/D conversion, it would be necessary to take a digital output from each system and to monitor the comparison via a single high quality D/A unit used as a reference. Perhaps this comparison could be done in turn via a number of possible D/A stages, such as those used in domestic CD players.

EDITING

Three primary means of editing digital recordings are available to the user today: splice-editing, electronic assembly editing, and random access editing. Although it is not intended to make this a comprehensive discussion of these methods, some operational differences are relevant.

Splice editing

As already mentioned, both the stationary head formats were designed specifically with the aim of allowing splice editing, and much time has been spent by the manufacturers over the last few years in coping with this requirement. Most people agree that a tape speed of at least 15ips is necessary for easy editing, and this is what has been adopted by both formats, although a lower speed of 7.5 ips had been initially adopted for the DASH format, implemented in Sony's PCM-3102 recorder and considered for the Prodigi format by Mitsubishi. It is possible that the 7.5ips speed will be used in cases where editing is not a primary requirement, and the Sony PCM-3402 has been designed with switchable speeds. Studer's DASH machine, and also Mitsubishi's Prodigi machine have one fixed speed of 15ips.

Experience of splice-editing with digital tapes is not wide as yet, but it is clear that considerably greater care is required than one would expect to use when editing analogue tape. Although this is not surprising, it slows down the editing procedure and requires that the edit is made correctly first time round, as the chances of getting a noise-free join once the edit has been picked apart, moved, and then rejoined are slim. This is not to say that the operation is impossible, just that the chances of success are limited. My own experience is that good splices can be made on both DASH and Prodigi machines as long as the correct thin splicing tape is used (in order not to affect head wrap excessively at the edit point), as long as the splicing tape is cut to be slightly narrower than the width of the tape itself, and as long as a tiny gap is left at the splice point. Still, it would be a nerve-wracking business with valuable master tapes.

The sound which one hears when rocking the reels of a digital recorder by hand is that of two auxiliary cue tracks which tend to be of low audio quality. These tracks are analogue recordings on Sony machines, but are pulse-width modulated on Studer and Mitsubishi machines so that the same heads can be used for recording the cue tracks as are used for the digital tracks. Cue track distortion is audibly worse on the X-86 than on the PCM-3402's analogue tracks, being only just adequate for editing purposes, and the sound has the annoying tendency to mute at very low tape speeds which makes location of the edit points difficult.

A final point that should be made is that the crossfade between old and new material at a digital splice is not controlled by the angle of the join (as it would have been with analogue tape), but by the crossfade generator within the machine. Machines often allow this time to be varied (the X-86 allows times up to 87 milliseconds, for example), but it is important to remember that this will apply to every edit that is made, whereas electronic editors like Sony's DAE-1100 allow the crossfade to be different for every edit, so as to make the best audible join for the material in question.

<u>Electronic editing</u>

If splice editing is not acceptable then the alternative is the expense of at least two tape machines, and sometimes an external controller or processor to perform the edit electronically. All the professional formats allow for this, and this is one of the main features which distinguishes them from semi-pro and domestic formats, although great efforts have been made to provide editing facilities for the cheaper EIAJ-format systems, with some success.

It has to be borne in mind that the assembly form of electronic editing relies on the control of at least two tape machines to create a master tape on which the edits, once performed, are fixed. Any alteration of edit points within the material requires re-copying of the entire programme, although there is the advantage that edits may be tried as many times as necessary before the result is comitted to tape.

As already pointed out, video-based systems like the PCM-1610 can only be dropped into record mode at frame boundaries, and this may not be accurate enough for editing as one frame of NTSC video is around thirty-three milliseconds long (or around half an inch of analogue tape at 15ips). For this reason, the editing processor has the job of finely locating the edit point, and of executing a crossfade within the frame in question, re-recording old material out of memory from the start of the frame until the edit point.

Stationary-head recordings may also be edited electronically, either using an external controller in the case of the X-86 Prodigi machine, or by linking two machines together directly in the case of the PCM-3402 DASH machine, which has an editing controller and synchroniser built in. The edit controller has to rely upon the ability of the stationary-head machine to perform a suitable crossfade at the join, thus this will be constant for all joins.

DAT has yet to prove its capacity for allowing editing, although this was initially a feature claimed by Sony for the format. None of the professional machines yet announced incorporate editing facilities, and if this never happens it would appear that DAT will remain a semi-pro medium along the lines of the EIAJ format. My own experience with DAT machines indicates that good glitch-free joins can be made by PLAY-PAUSE-RECORD editing (as one might try with a video recorder and a PCM-F1), which possibly indicates that the problem might not be too great. Certainly it was clear that a better join could be made using DAT in this way than could be achieved with the EIAJ format which tends to produce spits at a join.

Random-access editing

This area deserves brief mention here, although real products are only in their infancy. By storing digital audio in a random-access memory like a hard-disk store or semiconductor RAM, one can gain access to any section of the store in a very short time. Apply such principles to editing and you will have speeded up the process considerably, as one of the great time-wasters when editing is waiting for the tape to wind long distances. Products are beginning to appear which provide various lengths of random-access storage with control over the joining of pre-defined sections, often giving graphic display of the process, of audio waveforms, and so on. The advantages of such systems can be readily appreciated when it is seen that sections of material can be used many times over, and that a number of possible take lists can be compiled, which only require that the material is taken from the store in a different order.

The best digital tape recording medium may not be the best medium for editing, a fact which is clearly borne out in experience. This perhaps points to the potential success of random-access editing which would allow one to use a cheap but excellent recording format such as DAT, and a reasonably priced editor based on computer technology. As memory costs continue to drop, so will the cost of such systems.

DIGITAL INTERFACING

In the all-digital studio signals will pass between systems without analogue conversion. With the existence of so many recording formats this is already important for the transfer of material from one to another, and a number of methods exist.

AES/EBU interface

In an attempt at standardisation, the AES formulated a standard interface for the transfer of stereo digital audio data between equipment. This has now been adopted by most professional manufacturers, and is usually available as an option for their machines. It allows for the transfer of two channels of digital sound to and from devices at a number of sampling rates, without the need for an extra word-clock sync connection, as synchronisation is inherent in the interface.

As well as the professional interface, it is interesting to see that there is a domestic form of the same which appears on the back of some CD players, DAT recorders, and the Sony PCM-601 EIAJ processor. Although this uses a phono connector, it is the author's experience that connection can be made between such machines and other devices which use the professional AES interface such as the Sony PCM-1630, allowing successful digital transfer of material from EIAJ format to 1630 for CD mastering. This requires that the same sampling rate is employed in both devices, and that the receiving machine is capable of accepting the "domestic" flag in the transferred data. Some devices ignore this flag, some can be set to accept it or not, while others will refuse to accept data from a domestic device.

SDIF-2

The other common digital interface is a Sony standard, known as SDIF-2, which usually resides on BNC connectors, one per channel. This is used on most Sony equipment, and is provided (along with AES/EBU and Sony/Philips) on the PCM-2500 DAT machine. SDIF-2 is the interface used by Audio and Design in their digital interface modification for the PCM-701 EIAJ processor, and has been used with success over the last few years for the transfer of material from EIAJ-format to 1610.

Product-specific interfaces

Machines such as the X-86 Prodigi also sport a dedicated "digital copying" connector which is only for use in copying audio between certain machines from the same manufacturer, and does not conform to a recognised standard. Mitsubishi recommends the use of this connector for copying old X-80 tapes onto the new format.

SUMMARY

This paper has discussed a number of operational questions concerning the use of two-channel digital audio recorders in practical environments. It has pointed out that, although there are many formats, no one system has succeeded in fulfilling all the requirements of the user.