CD SIGNAL PROCESSING

PART TWO BY VIVIAN CAPEL

THE WISDOM OF REED AND SOLOMON

The binary wordplay needed to squeeze sixty minutes of music onto a three inch disc includes symbolising, interleaving of audio sample frames, three-zero humps, discardable merging bits, delay lines, triple parity checks and double decoding. They say it's great for Heavy Metal...

A lot has been written as to how cd signals are recorded, but some of that published elsewhere has been vague and often inaccurate. In our last article we saw how the player tracks the disc and produces a data stream; now we shall see how the signal is processed, how a quart is compressed into a pint pot, and what happens when dirt or disc imperfections cause signal errors. We hope also to clear up a few misconceptions along the way.

Many readers will be familiar with the principles of digital recording, but for those that aren't we will start with the basics. When a signal in a recording medium, whether an electric voltage or current, magnetic field, or record groove, varies in proportion to the sound pressure wave that caused it, it is termed an *analogue* signal. It is analogous to the original sound.

Photograph of the TR8848 cd radio cassette recorder by courtesy of Pye

Now it is not easy to get the signal to correspond exactly. The system may discriminate in favour of, or against certain frequencies, components can resonate producing frequency peaks, spurious harmonics can be added, and so can noise, frequencies can interact generating sum-and-difference products that were not in the original. Any or all of these can occur during recording and again at playback. The end result is *distortion*.

DIGITAL RECORDING

An alternative to attempting to accurately record the actual wave is to measure it, record the measurements, then reconstitute the wave from those measurements. It is obvious that once the measurements have been taken nothing can happen to degrade the signal. Noise and distortion can have no effect as long as the reproducer can recognise the numbers. The only problem that can arise is if some of the measurements are lost.

Distortion could arise if the original measurements were inaccurate or they were inadequate to truly represent the analogue signal. There must therefore be a sufficient number of readings or samples so that the waveform can be properly defined. According to the *Nyquist Theory*, the sampling rate must be at least twice that of the highest frequency to be represented. In the case of the compact disc, the sampling rate is 44.1kHz, which is sufficient to describe a sine wave of 20kHz, the highest audio frequency recorded. The top frequency is bound to be a sine wave because if it was not, the presence of still higher harmonics would be indicated.

Should frequencies be present that are higher than half the sampling rate, the samples appear to the decoder to be values of a much lower frequency, (see Fig.1). This effect is known as *aliasing*. To prevent it, all frequencies above the required range must be severely curtailed. This is done by means of a steep top-cut filter. Having the sampling rate slightly higher than twice the upper frequency limit, besides giving a safety margin, also allows for the filter roll-off.

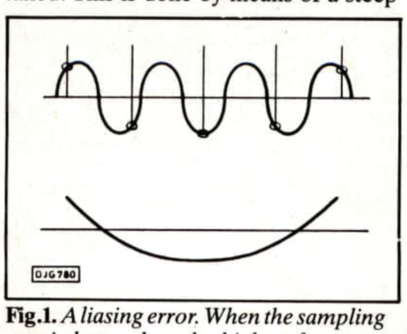

Fig.1. *A liasing error. When the sampling rate is lower than the highest frequency, spurious low frequencies are produced*

QUANTISATION

Each sample is represented by a whole number, so the larger the number of measurable levels the more accurate the sample. For example, if temperature is to be measured in whole degrees, the Fahrenheit scale having 180° between freezing and boiling points will give a more accurate result than Celsius which has only 100°. The error caused by a value falling between two levels is termed *quantisation distortion*. The percentage of error caused say by 2.5 being measured as 2, is far greater than 200.5 being rounded off to 200. It thus follows that the effect of quantisation error is much greater at low signal levels. Decaying sounds die away in a series of steps rather than smoothly, producing what can best be described as a 'crumbling' effect.

One way of overcoming this is to space the levels closer together at the low end, a method known as *non-linear quantisation*. With the compact disc 65,535 levels plus 0 are measured, giving 65,536 possible numbers. This is sufficiently high to give very low quantisation distortion even at low levels, so non-linear quantisation is not necessary.

BINARY CODE

If the sample was recorded as a decimal number, ten different states would be required to represent each digit. This would produce the same problems as an analogue signal, with slight errors producing totally different numbers, especially if these occurred in the left-hand digits. With the binary code, only two states are required to represent 0 and 1, so a series of pulses and gaps can be used. The method is thus termed pulse-code modulation, and is far less likely to result in ambiguity than the use of a decimal number.

While many readers will be familiar with the binary code we will ask them to bear with us for a while for the sake of those who are not. In the binary code, each 1 starting from the right indicates a power of 2 starting with 2^0 which is 1, followed by 2^1 which is 2, then 2^2 which is 4, 2^3 which is 8, and so on. The appearance of a 0 in any position indicates the absence of that power from the complete number. Any number up to 65,535 can be represented by 16 binary digits which is therefore the number of digits or bits used for the compact disc.

For comparison, 15 digits can represent up to 32,767 decimal, 14 digits, up to 16,383, and 13 digits up to 8,191. So the extra few bits mean a lot! Actually, 16 bits is somewhat a case of overkill as the excellent quality achieved with the BBC fm stereo broadcasts is obtained after transmission along studio link lines at 13 bits. However, the excess is undoubtedly to wisely achieve a higher-than-necessary standard from the beginning, and so avoid the need for upgrading later when advances in associated equipment could render a lesser specification obsolete.

FRAME FORMAT

As we have seen, the 65,535 possible signal levels are represented by a 16 digit binary number which is called a **word**. The rh and lh stereo channels are recorded consecutively and so consist of alternate words. This is possible because the data is recorded much faster than it is sampled. Samples are taken at the rate of 44,100 per second from each channel which is 88,200 samples per second. If these were recorded as 16-digit words per sample, this would be 1,411,200 digits or bits per second. The recording frequency is 4,321,800 bits per second, so both channels as well as other data can be recorded consecutively without running out of time. In the reproducer, the samples are stored in a memory as they come off the disc and are then clocked out at the required rate.

Actually, the samples are not recorded as straight 16-bit words. The units need to be shorter because each one is nearly

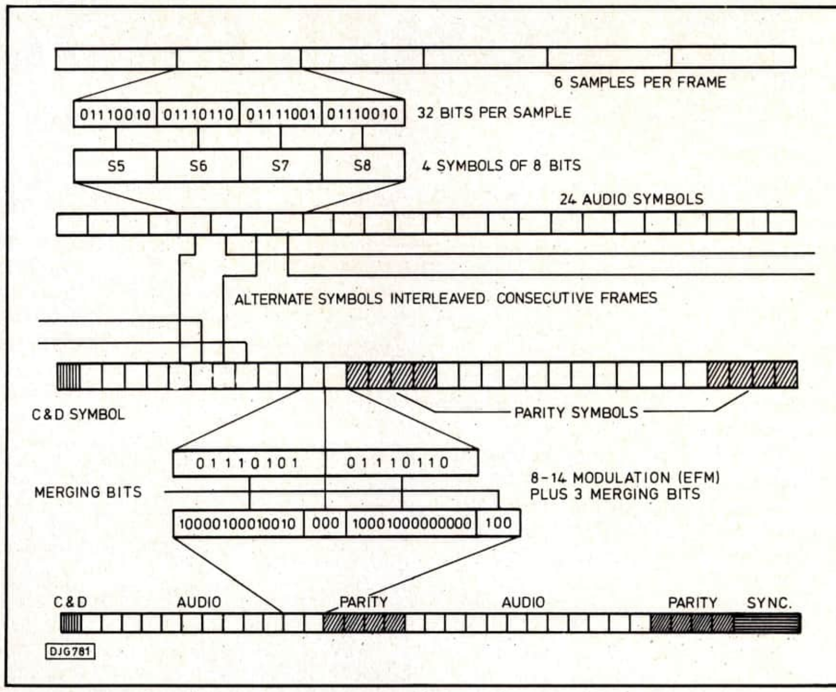

Fig.2. *A complete frame showing composition and successive processing*

doubled in length later, for a reason we shall see. This would produce unwieldly units which the digital circuits would find difficult to handle. So each word is split into two 8-bit sections called *symbols*.

The two stereo channels are sampled simultaneously so they constitute a single sampling of the sound field at any instant. Hence both are referred to as a single *sample* in compact disc terminology.

There is a set sequence of audio symbols, parity symbols, control and display symbols, and synchronising bits. A complete sequence is known as a *frame*, all frames having the same sequence. (Fig.2.)

The actual sequence is: first, one *control and display* symbol, which gives data as to title and composer, timing and track number and any other required data. This information can then be displayed on a read-out panel in the player, and the player's computer control can search for a particular track number when so instructed. Finding a track is thus much easier than the hazardous pickup setting down procedure which has to be followed with the standard lp record. Although only one 8-bit symbol is allocated per frame, there are 7,350 frames per second, thus giving 58,800 data bits per second which is ample to record the required control and display information.

Second in the frame come twelve audio symbols which represent six 16-bit words or three complete stereo samples. Next, there are four parity symbols whose function we shall discuss later. These are followed by another twelve audio symbols giving a further three audio samples, after which are another four parity symbols. Finally, there are 27 bits of synchronising data which the decoding circuits recognise, and so can tell where one frame ends and the next one starts. As these come between each frame they can be considered either as ending a frame or starting the next one.

A complete frame thus consists of 33 symbols of eight bits, which include six audio samples, plus the 27 sync bits. As there are 7,350 frames per second, this gives us the $6 \times 7,350 = 44,100$ samples per second that we have previously noted.

Now, the audio samples are not recorded in succession, they are interleaved with those of an adjacent frame by a system of delay lines having a delay time of one frame. Thus frame A is delayed to be concurrent with frame B and the symbols are taken from each in turn. Thus the even ones from A alternate with the odd ones from B. The evens from B are meanwhile delayed so that they in turn are alternated with the odds from C whose even symbols are alternated with D and so on. The interleaving process is thus not just between a pair of adjacent frames, but involves those on either side. It is therefore known as the *Cross-Interleaved Reed-Solomon Code*.

In the playback decoder, the process is reversed, with alternate frames being again applied to delay lines and symbols from each being interleaved, only this time they are really being de-interleaved as the process restores the original positions.

The obvious question is, why go to all that trouble? The answer is that because errors and drop-outs are usually introduced by blobs or scratches on the disc surface, they tend to come in groups rather than being spread out. If frames were recorded successively, a whole frame or even more could be lost irretrievably. By interleaving them the loss is spread out to parts of several frames instead of one complete one. This enables the error correction circuitry to deal with the errors by averaging from adjacent symbols and other means that we shall discuss presently. It is rather like putting your money on several horses rather than all on one, you are less likely to lose the lot!

SPOT LIMIT

The number of channel bits that can be recorded depends on the diameter of the light spot that scans the disc. It does not have a sharp boundary so the rated diameter is that at which the intensity is at half value of the centre illumination. It is given by:

$$d = \frac{0.6\lambda}{NA}$$

in which d is the spot diameter, λ is the laser light wavelength, and NA is the numerical aperture of the objective lens. For a wavelength of 0.8 μm and a NA of 0.45, the diameter is 1 μm.

A smaller spot size could be obtained by increasing the NA, but this reduces the field of focus and would pare down manufacturing tolerances that are already fantastically small. The focusing tolerance for example is just $\pm$0.5 μm, less than the wavelength of the light used!

EIGHT-TO-FOURTEEN MODULATION

It follows from the foregoing that there is a limit to the amount of information that can be crammed onto a disc. However, a rather ingenious method is used to increase it although at first glance it would appear to make matters worse.

As is well known, the recording takes the form of a spiral track of pits that are excavated at the back of the disc, after which the rear surface is silvered. The track is read from the front through the transparent disc material, so it appears to the light beam as a track of humps.

The humps are 0.12 μm high, so light reflected from them is shifted by a quarter wavelength compared to that reflected from the surrounding flat areas, which are termed *land*. Interference thus causes darkening as the hump passes, and the effect is increased by diffraction which scatters some of the light, because the hump width of 0.6 μm is smaller than the light wavelength.

A common misconception is that the humps and spaces denote the 1s and 0s of the binary data respectively. This is not so; a 1 is indicated by either a leading or lagging *hump edge*. Such an edge represents a transition from one state to another, light to dark or vice-versa, and it is these changes that the detecting circuits interpret as a 1. What follows whether light or dark, is interpreted as a 0. If the space between two edges, whether hump or land, is twice the unit length, the detector reads it as two 0s, if three times the length, three 0s and so on until the next edge.

A hump three units long thus conveys five digits, 10001. So by adopting this method, more information can be recorded than if each hump was a 1. But this is not the only means used to try and get a quart into a pint pot.

It is obvious that as the spot cannot read two edges at the same time, the hump cannot be shorter than the diameter of the spot. So in the case of a number having different alternate digits (010101...) the unit of space occupied by one digit cannot be made shorter than the spot size. However, if all numbers had say three or more adjacent 0s (00010001; 10001000;...) the space unit allocated to each digit could be shorter, and more information could be packed in.

With the standard binary code of course, this is not so; there are many numbers that have different alternate digits. So although the standard binary code has been used in recording and processing up to this stage before actually modulating the recording laser it is converted to a different code altogether. Tables 1 and 2.

This code has no logical progression like the binary code; it is an arbitrary collection of digital numbers, each being assigned to represent a standard binary number. The important thing about them is that each one has three or more adjacent 0s. A 'dictionary' of logic gates

Decimal	Binary	Decimal	Binary
1	1	60	111100
2	10	64	1000000
3	11	70	1000110
4	100	80	1010000
5	101	90	1011010
6	110	100	1100100
7	111	110	1101110
8	1000	120	1111000
9	1001	128	10000000
10	1010	130	10000010
11	1011	140	10001100
12	1100	150	10010110
13	1101	160	10100000
14	1110	170	10101010
15	1111	180	10110100
16	10000	190	10111110
17	10001	200	11001000
18	10010	256	100000000
19	10011	300	100101100
20	10100	400	110010000
21	10101	500	111110100
22	10110	512	1000000000
23	10111	600	1001011000
24	11000	700	1010111100
25	11001	800	1100100000
26	11010	900	1110000100
27	11011	1,000	1111101000
28	11100	1,024	10000000000
29	11101	1,200	10010110000
30	11110	1,400	10101111000
31	11111	1,600	11001000000
32	100000	1,800	11100001000
33	100001	2,000	11111010000
34	100010	2,048	100000000000
35	100011	2,500	100111000100
36	100100	3,000	101110111000
37	100101	3,500	110110101100
38	100110	4,000	111110100000
39	100111	4,096	1000000000000
40	101000	4,500	1000110010100
41	101001	5,000	1001110001000
42	101010	6,000	1011101110000
43	101011	7,000	1101101011000
44	101100	8.000	1111101000000
45	101101	8,192	10000000000000
46	101110	9,000	10001100101000
47	101111	10,000	10011100010000
48	110000	15,000	11101010011000
49	110001	25,000	110000110101000
50	110010	60,000	1110101001100000

Table 1. Decimal to binary conversion chart. *It is evident from the number of consecutive 1s, or where only one or two 0s are consecutive, that to substitute a 16-digit binary number with one containing no consecutive 1s and no fewer than three consecutive 0s, a much larger number of digits is required, no fewer than 28 in fact. To avoid excessively long words and unnecessarily complex processing circuits, in the cd system the 16-bit words are split into two 8-bit symbols which are changed to 14-bit symbols. By permitting digits to be recorded in a third of the otherwise required space, this* eight/fourteen modulation *saves 25% of disc space.*

digits	decimal max	digits	decimal max
1	1	9	511
2	3	10	1,023
3	7	11	2,047
4	15	12	4,095
5	31	13	8,191
6	63	14	16,383
7	127	15	32,767
8	255	16	65,535

Table 2 Maximum Value: *The maximum value for a given number of digits is when all of them are 1s. The maximum for up to 16 digits are:*

is used to accomplish the conversion in the recording circuit, and a similar one in the player converts the numbers back to binary code for decoding. The numbers are chosen to best suit the architecture of the logic gate chip, but their actual composition doesn't really matter as long as the recording and playback dictionaries are the same.

It follows then, that the unit space for each 0 can be shorter than the diameter of the light spot because it will never encounter a number with less than three 0s. It is in fact 0.3 μm, a third of the light spot diameter. There is a snag though (isn't there always?). You cannot represent all the binary numbers having 8 digits, with another set of 8 digit numbers having no fewer than three adjacent 0s and no adjacent 1s. The new set must be longer. It needs, in fact, no less than 14 digits to represent each 8-digit symbol (hence the designation: eight-to-fourteen modulation). This incidentally is why the original 16-bit words were split into 8-bit symbols; they would otherwise have grown to an unwieldy 28 bits, and the dictionaries for conversion would have been much larger and more complex. Naturally, this increase to 14 bits erodes some of the advantage gained by using shorter space units, and there is a further problem.

There must as we have seen be a minimum of three consecutive 0s, but there is also a maximum. When there is a long run of consecutive 0s there are no light transitions and so there is no signal change. The circuits that count each unit and so determine the number of 0s could thus get out of synchronisation because the pulses provided by the 1s which they latch on to, are absent. The effect is like a tv receiver in which the frame sync pulses are missing due to a fault. The picture remains steady for a second or two but then slowly drifts up or down the screen and starts rolling.

Another effect of a long run of 0s is that the light beam modulation frequency drops to a low value. The servo systems that control the tracking and automatic focusing operate at low frequencies (below 20kHz), so they could be confused by any modulation within their frequency range and thus misfunction.

Because of these effects, the maximum number of consecutive 0s that can be permitted is ten. This is not difficult to arrange when designing the new 14-bit numbers, but a snag arises if a number that concludes with a string of say eight 0s is followed by another that starts with seven. Now we have a string of fifteen consecutive 0s and the maximum has been exceeded. There are quite a few combinations that could exceed the limit of ten. Another difficult situation would be if a number that ended with a 1 or 10 was followed by one that started with a 1. In this case there would not be the minimum of three 0s between the 1s so the spot would be too large to read both 1s.

To overcome these problems three extra bits are added at the end of each symbol. These convey no information and are discarded when decoded in the player, but their function is to prevent consecutive or near consecutive 1s and to break up successive 0s between two following numbers. Their composition depends on whether they are separating 1s or 0s. If more than one combination satisfies the requirements, the one chosen is that which gives the lowest ratio between hump and land length. This ratio is known as the *digital sum value*, and the lower it is the lower the noise in the servo frequency band. The extra three bits are known as *merging bits*.

So the addition of the three merging bits to the 14-bit word extends the original 8-bit symbol to 17 bits, seemingly not a very good exercise in economy and space saving! However, because of the recorded unit being only a third of its otherwise minimum length, there is still an overall improvement in information density of 25%. Without it, the maximum time per disc would be 48 minutes instead of 60. Thus many musical works can be recorded complete on one side of a disc that otherwise could not.

RECAP

At this point perhaps we can pause a little to digest what we have already discussed. The various processes are illustrated in Fig.2 so we will use it to recap. We note at the top that there are six stereo samples in each frame, each consisting of one 16-bit word for each channel, making a total of $2 \times 16 = 32$ bits for the complete sample. These 16-bit words are each divided into two 8-bit symbols to facilitate later processing. Thus we have four symbols per sample, and 24 for each frame.

Next we see that alternate symbols are interleaved with those of adjacent frames on either side, so that errors or drop-outs are spread over several frames instead of concentrated in one or two. This makes them easier to compensate for later.

Eight parity symbols are added in two blocks of four, (of which more in the next section), and also one control and display symbol at the beginning. Next, the 8-bit symbols are converted into 14-bit symbols that avoid consecutive 1s, and enable the space occupied by one digit to be made only a third of the laser spot size. To prevent consecutive 1s between adjacent symbols or an excessively long run of 0s, three merging bits are added at the end of each symbol which are discarded during replay. This brings the total bits per symbol to 17, yet in spite of the extra number an overall space saving of 25% is achieved.

ERROR CORRECTION

A major consideration is the possible loss or corruption of the digital signal due to imperfections in the disc, which considering the microscopic size and the astronomical number of humps in each disc is virtually inevitable. The effect of dirt on the disc surface, though minimised by being out of focus, can also cause loss of signal.

The system of error correction built in to the compact disc is therefore quite elaborate. Firstly, consider how simple corrections of a digital signal can be made by the use of *parity bits*. Imagine a data stream of 8-bit words. Some words will have an odd number of 1s while others will have an even number.

For example, 01101011 has an odd number with five 1s, and so has 01000101 with three. But 01100110 has four, an even number, just as has 01110111 with six. Let us now add an extra bit to each of these so that they always have an even number of 1s. The first number needs an extra 1 and so becomes 01101011<u>1</u>; the second also needs another 1 and thus becomes 01000101<u>1</u>. In the case of the third number it already has an even number of 1s, but as all words must be of the same length it must have an extra digit, so we add a 0. It thereby becomes 01100110<u>0</u>. This applies also to the fourth number which the added 0 turns into 01110111<u>0</u>. We have underlined the added digits which we call parity bits because they put all the words on a par respecting the number of 1s they contain.

Later, after recording the stream, we can get a circuit to check each word for an even number of 1s, and if they are correct, to then ignore every ninth digit which is the meaningless parity bit. If though it finds a word with an odd number of 1s, it knows that the word is incorrect. Of course that doesn't tell us which bits are wrong or how many bits in the word are faulty, but it does identify an inaccurate word.

If the stream represents an analogue audio signal, the faulty word can be erased and its value calculated by taking an average of the adjacent words. The value may not be 100% accurate, but it will not be far out and most likely far closer to the original than the erased faulty word, especially if the defect was with one of the high value digits at the beginning of the binary number. This method of correction, known as *interpolation*, is one of several used with the compact disc, but it is possible to identify and correct the actual faulty digit within the word and thereby reproduce the word correctly.

Such identification can be made by forming the digits into parity blocks. This can best be understood by reference to Fig.3. Here we have for the sake of simplicity three 4-digit words, x1–x4; x5–x8; and x9–x12 arranged in a block. Each has a parity bit added at its end just as before, these being p1, p2, and p3.

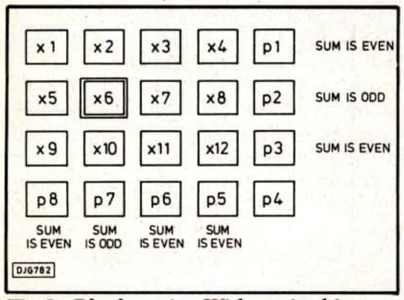

Fig.3. *Block parity. With parity bits added to rows and columns, an odd sum for a row and column identifies the incorrect bit at their intersection*

In addition to these there are parity bits added to the bottom of each column, p8, p7, p6 and p5, with one at the bottom of the column of line parity bits, p4.

The parity bits at the end of each line make each line have an even number of 1s, while those at the bottom of the columns make the columns even. Let us say now that data bit x6 is wrong; this will make the sum of the 1s in that line odd instead of even, and also it will make the sum at the bottom of that column odd. So by identifying the incorrect column as well as the line, the digit at the intersection is positively identified as the culprit. As binary digits have only two values, changing the faulty one to its opposite must make it right. The bit p4 seems redundant, but it serves to check on the other line parity bits as one of those could be faulty just as much as a data bit.

A question may be asked here, how do you convey a two-word dimensional block along a one-dimensional data stream? The answer is like a tv picture, line by line. It is only drawn and described as a block so that we humans can see how it works. The decoder puts the sequence in a memory then sees it as a number of sums; it adds the digits after every fifth bit (end of line), and it also sums the first, fifth and ninth digits, the second, sixth and tenth, the third, seventh and eleventh, and the fourth, eighth and twelfth. When it finds a couple of errors it calculates the intersection and promptly changes the digit there. In the compact disc, blocks are formed from groups of symbols rather than individual digits. This achieves the maximum amount of error correction with the minimum number of parity symbols.

The number of incorrect symbols in a block that can be both identified and corrected in the decoder is limited. However, if the faulty symbols are identified and labelled first, the decoder can correct twice as many.

DECODING AND CORRECTION

As the data stream comes in from the photodiodes, the start of each frame is identified by the synchronising pulses. Next, the 14-bit symbols are converted back to 8-bits by the logic 'dictionary', and the merging bits discarded.

Now the frames are fed into a buffer memory and clocked out at the other end by a quartz controlled clock. The input rate may vary due to wow or other causes, but the output rate is constant so wow can have no effect on the reproduction. The buffer is kept nominally 50% full to allow regulation in both directions, but the actual amount it contains is an indication of whether the motor is running fast or slow. The amount contained is therefore used to generate a control signal for the motor.

The control and display symbol, which is the first of the 33 in each frame is removed and routed to its own circuit. The remaining 32 are then fed in parallel to 32 inputs on the first *of two* decoders. Alternate inputs have delay lines having a delay length of one symbol which thereby restore the interleaved symbols from successive frames to their correct position. The decoder identifies any faulty symbols and corrects one. Two could be corrected here but with a lower degree of accuracy. So if there are more than one, all are given labels and are passed on uncorrected. The parity symbols serve their purpose in the identification and so go no further.

The audio symbols are dispatched to the second decoder via delay lines that space successive samples by 8-frame intervals so that error groups are well spread out and thereby diluted. As the faulty symbols have been pre-labelled by the first decoder, more can be corrected here, up to four in fact. If there are more than that, all are passed without correction but still retaining their labels.

INTERPOLATION

Up to four symbols in each frame can be corrected; beyond this interpolation is necessary. Labelled symbols are erased and their value is estimated from adjacent ones. As odd and even rh and lh channel samples within each frame are interleaved, up to seven consecutive samples can be interpolated. However, owing to the 8-frame spacing introduced by the second decoder, this is multiplied across 56 frames. A two-frame delay line enables a complete frame to be substituted if required, but this reduces the 56-frame capability by two, and delay lines in the recording circuits reduce it by five or six. So reliable interpolation can take place over 48 frames.

This corresponds to some 12,300 data bits or 7.7mm of track length. Distortion increases over the interpolated portion and the bandwidth drops from 20kHz to 10kHz. The maximum burst of errors that can actually be corrected is 4,000 data bits which corresponds to 2.5mm of track length on the disc.

MUTING

Correction and interpolation will cope with the majority of defects and small blobs of dirt. If though a stream of uncorrectable errors occurs that exceeds the capability of the system, the last resort is muting. The gain is reduced so that the aural effect is that of a drop-out rather than a burst of noise which a stream of uncorrected errors would produce. The gain does not drop suddenly as this would produce harmonics and a sharp click. Instead the drop and rise follow a cosine curve. The drop starts from 32 samples before the error stream, and rises 32 samples after it.

SMOOTHING THE STEPS

Following correction, the alternate rh and lh signals are separated and applied via filters to the two d-a converters. This though is not the end of the story. After conversion the analogue wave is in the form of steps at the original sampling frequency of 44.1kHz. These must be smoothed out to reproduce the original waveform, and to remove the harmonics of that frequency which could beat with others such as tape bias.

To cut frequencies of 44.1kHz and upwards, a filter of at least 50dB attenuation is required, yet this must have no effect on frequencies of 20kHz which is little more than an octave below it. This is a tall order and one of the principal headaches for the designer. Complex filters having the desired characteristics are likely to produce ringing or other spurious effects in the audio passband.

OVERSAMPLING

One method commonly adopted for overcoming the filter problem is oversampling. With this, each sample fed to the d-a converters is electronically remembered and repeated before the next arrives. With twice the number of samples being presented to the converters, the frequency of the samples is obviously doubled, and the process is termed *two-times oversampling*. In many players the signal is repeated four times so giving *four-times oversampling*. The sample frequencies are thereby increased to twice and four times respectively that of the original sampling rate, that is 88.2 and 176.4kHz. The gap between these and the highest audio frequency is much greater than that between the original sampling rate and the audio band, so the filter can have a less steep slope, be less complex, and have less of an effect on the audio signal.

Of course, the number of samples actually recorded per second has not changed, only the fact that each is used four times (in the case of four times oversampling) in the conversion back to analogue. As it stands, this gives four successive identical samples, which means a level step at a quarter of the frequency, in other words we are back to square one with 44.1kHz steps and nothing has really changed.

Oversampling is therefore only effective if the intermediate steps that follow the first of a new sample are varied. One method of doing this as used by Philips is by means of what is called a transversal filter. (Fig.4.) This consists of 24 delay lines connected in series, each delaying the signal by one sampling period, so giving a total delay of 24 samples at the end. Part of the signal is tapped off at each line four times during each sample period, and it is multiplied by a constantly changing coefficient, then passed to an adding circuit. The products of sample and coefficient are thus added to each of the respective quarters of the preceeding 24 signal samples.

The coefficient is a 12-bit word which is added to the 16-bit sample to make a 28-bit word (adding two binary numbers in this manner multiplies them). Its value is chosen for each sample so that the summation does not introduce any extra bits. By this means three intermediate values between the signal samples are obtained.

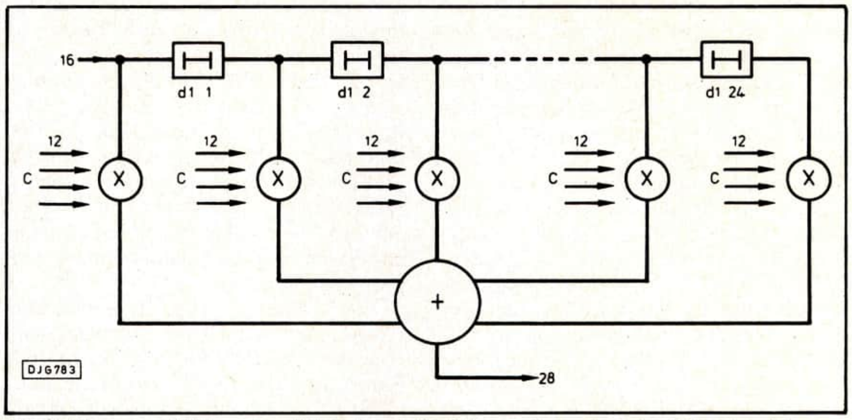

Fig.4. *Transversal filter, signal is delayed for 24 sample periods by 24 series delay lines. Output from each is multiplied 4 times by 12-word constant, and added*

14 OR 16 BITS?

Early 16-bit decoders were quite inaccurate often tending to ignore or falsify the least significant couple of bits. Thus their resolution was no better than a 14-bit decoder, in fact they could be less accurate because their errors were not consistent. Another factor was that 14-bit devices operated much faster than the 16-bit. This permitted the use of oversampling which requires fast operating circuits to handle the high sampling frequencies. Some makers, notably Philips, therefore preferred to use an accurate 14-bit decoder with oversampling.

Quantisation noise is produced in the sampled frequency band which is proportional to the size of the steps. If the step size is doubled by reducing the bit number by one, the noise level is also doubled, so increasing it by 6dB. Thus a basic 14-bit system has 12dB more noise than a comparable 16-bit one, actual values being about −84dB and −96dB respectively.

However, with four times oversampling, the noise is spread over a wider frequency spectrum, so that falling within the audio range is some 6dB less, bringing the noise down to −90dB.

When rounding off to 14 bits, there is a rounding off error which at low frequencies is almost the same as adjacent samples. At a sampling rate of 176.4kHz, most of the audio spectrum is at a comparatively low frequency. This error is stored for one sample period and added to the next sample in antiphase, thus virtually cancelling it out.

The effect is to reduce the distortion introduced by casting off two bits, and also reduce noise by some 7dB. Total noise thus becomes −97dB, marginally better than that of the 16-bit system. Of course noise levels of this order are swamped by the music signal and even by the ambience of the recording studio, so are really of academic interest only. However, it does show that by a little circuit ingenuity, 14-bit players can equal the performance of some using the full 16 bits. However, time marches on and so does technical development. 16-bit decoders have improved since those early ones and it should be expected that they will eventually become standard along with oversampling.

It may be thought that all this complicated messing around with the signal, chopping it up, delaying parts of it, interleaving it, sticking in other bits and pieces, would surely produce a marked deterioration of the result. Well it certainly would if the signal was analogue, but this is the beauty of a digital system. It just records numbers that instruct the dna converter what to produce. You can play about with those numbers and record them any way you like providing they are assembled correctly at the end.

When considering these parameters though it must be remembered that there is a degree of overkill. The excellent BBC music fm transmissions use digital studio to transmitter links that employ only 13 bits and have a high frequency limit of 15kHz. As the upper limit of human hearing is 16kHz, and that only with young adults in their twenties, it can be seen that the cd system has been engineered to give better results than are really necessary.

Conditions in the recording studio, acoustics, microphone placements, and tape editing now have far greater effect on the reproduction than any factor associated with the cd player, other than an actual fault. The same is true for other parts of the domestic side of the system, especially the loudspeakers.

The stereo lp was in its day said by many to be the last word in domestic sound reproduction, so one has to be careful in making similar pronouncements for the cd. It is difficult though to see in what way it could be improved. But on reflection there is one thing. In spite of the mass production of discs and the millions sold, the cost remains far too high. PE

SEEDY CUSTOMER

Heroin-smuggling ring leader Paul Dye thought he was being clever in using a Psion Organiser to store temporary details of deals. Undercover customs officers secretly watched him and saw him using the Organiser on a drug run to Pakistan. When the arrested him they found the Organiser with $65,000 in cash. "It won't do you any good", he boasted, "I've erased everything". With crossed fingers, customs officers took the Organiser to Psion who worked on the eprom memory and recovered an almost complete record of the drug dealings. Had the Organiser used eeprom, like the Philips fts cd player, he might have got away with it. **Ed.**