PLAYING THE COMPACT DISC

BY VIVIAN CAPEL

BUMPS ON THE LANDSCAPE

The little audio discs don't rely only on high resolution and speedy revolution – they are played upside down and inside out, and have error checks on their error checks.

The shining silver discs with the crystal clear sound are becoming increasingly familiar. It remains a mystery to many though, just how the information is abstracted from the disc and in particular how the playback head manages to follow the recorded track, seeing that there is no physical contact between them. This is especially astonishing when the dimensions are considered. The track pitch is 1.6μm, and the track width a mere 0.6μm. In comparison the groove of an lp record is at its narrowest about 50μm wide.

The digital data is recorded as a spiral track of pits starting from the inside of the disc and ending near the outer edge. This is the opposite of the lp. Another opposite is that the disc is played on the underneath surface instead of the top, so to the pick up head it is travelling anti-clockwise.

Rotation speed of the lp is constant, which means that the modulations are cramped together at the centre where the groove circumference is small thereby increasing distortion, and expanded at the outer grooves thus wasting space. The compact disc rotates with a constant track velocity of 1.2 m/s, which means that the rotational speed is variable, slowing from 500 to 200 rpm approximately from start to finish.

The disc is made of plastic polycarbonate which is stamped or injection moulded to the required size and shape with the billions of modulation pits. The pitted surface is silvered in a mist of ionised aluminium which deposits a layer of some 0.04μm thick. Then the silvered surface is coated with a hard lacquer to seal and protect it. The modulations are read by a beam of light from a laser, but as it does so from the other side through the transparent disc material, the pits appear to it as humps.

All disc dimensions must be to extremely close tolerances. The thickness which is 1.2mm has a tolerance of ± 0.1mm; the disc flatness to 0.6° and the pit edge position to ± 0.05μm. The centre hole is 15mm in diameter but must be very accurately positioned. An eccentricity of 0.1mm would cause a beam deviation across 60 tracks! The disc has a small pilot hole drilled for the initial stages of manufacture, but the main hole is punched at the final stage. It is determined optically by a laser beam and is the exact centre of the track spiral, not necessarily of the disc perimeter. If after three attempts the laser cannot determine the centre, the disc is scrapped.

Photograph by courtesy of Morphy Richards CE Ltd.

DISC DIMENSIONS

Diameter	120mm
Thickness	1.2mm ± 0.1mm
Centre hole	15mm
Programme start radius	25mm
Programme finish radius (max)	116mm
Rotation	Anticlockwise to laser
Maximum recording time	60 minutes
Channel number maximum	4 with reduced playing time
Track pitch	1.6μm
Pit width	0.6μm
Pit or space length per digit	0.3μm
Minimum pit or space length	0.9μm
Pit depth	0.12μm±0.01μm

SCANNING BEAM

It is necessary that the scanning light beam focus down to a spot comparable to the dimensions of the humps otherwise more than one hump would appear in the spot at the same time and it may even spread over to an adjacent track. To get a light spot that small is difficult due to the effects of chromatic aberration. This is the appearance of coloured rings around an object due to the different angles of diffraction through a glass lens, of light of different wavelengths.

Achromatic lenses are made by combining elements of different types of glass so that the diffraction differences of each are cancelled, but these are very expensive. A further problem is that the

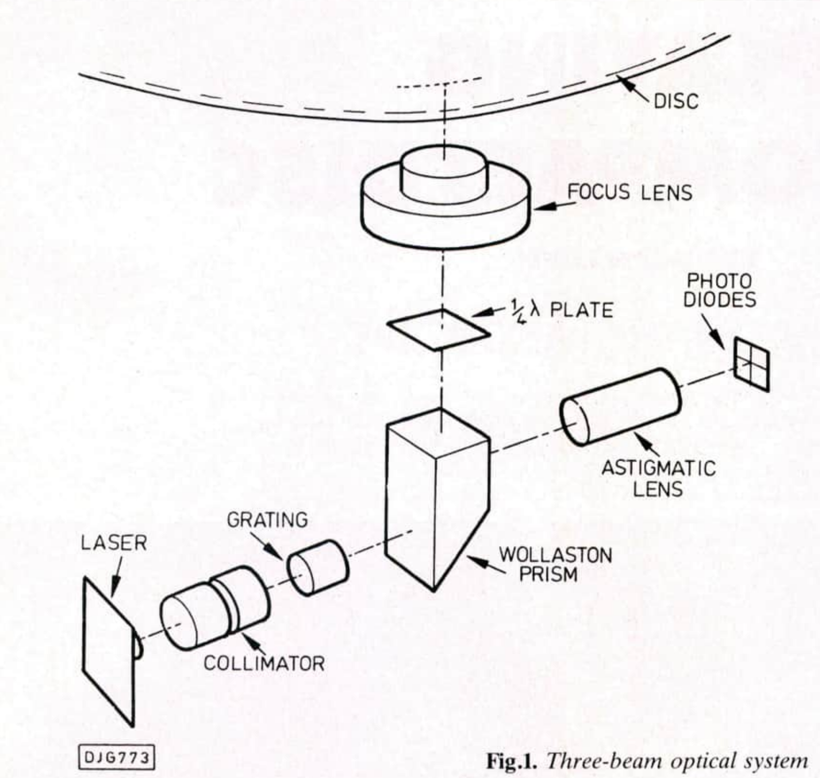

Fig.1. *Three-beam optical system*

humps must appear dark compared to the surrounding disc space (called *land* in CD terminology) so that they will effectively modulate the reflected light beam.

Both these problems are solved by using a laser. Having a single wavelength, there is no chromatic aberration and a relatively cheap lens can be used to produce a very fine spot. As laser light is coherent, that is all the waves are in step and in phase, light reflected from the hump which is raised about a quarter wavelength from the surrounding surface, is displaced half a wavelength compared to that from the surface. Cancellation occurs thus resulting in a darkening of the reflected beam. The hump width of 0.6μm is less than the laser wavelength of 0.8μm, so there is a degree of diffraction and light scattering which gives a further darkening effect.

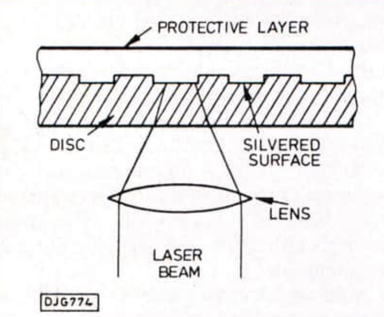

Fig.2. *Refraction of beam through transparent disc material*

Laser technology has improved considerably on the old gas lasers which were large and expensive. Those in current use are solid state, of aluminium gallium arsenide, and operate at low voltage which is another advantage. A further feature with some models is that the laser together with part of the optical system is contained in an easily replaceable plastic arm. However, replacements should be rare because life expectancy is now some 25,000 hours, which is quite a lot of playing time, nearly 3 years playing 24 hours every day!

THREE-BEAM SYSTEM

There are two basic optical systems, the three-beam and the single-beam. The operation of the single beam type can best be followed if we consider the three-beam one first. (Fig.1).

From the laser, the beam passes through a device called a collimator lens which ensures that the sides of the light beam are parallel and do not diverge. From there it travels through an optical grating which splits the beam into three. The light intensity of these is 50% for the main central one, and 25% each for the two side beams.

Next, the three beams are intercepted by a Wollaston prism which deflects them through a 90° angle and also performs a vital function on the return journey. From there, they pass through a quarter wavelength plate which rotates the plane of polarisation through 45°. This too is also necessary for the return pass.

Finally on the outward journey, comes the focusing lens which concentrates the beams into three tiny spots on the underside of the record.(Fig. 2). The point of focus is not at the disc surface through which it passes, but on the silvered layer above it. Thus the lower surface is out of focus which thereby reduces the effects of any surface blemishes on it.

After reflection from the silvered surface and modulation of the main beam by the humps it encounters, the beams pass back through the lens and again encounter the quarter wavelength plate. They are thus rotated a further 45° which means they are now 90° different from the forward going beams.

From there, the beams once more reach the Wollaston prism. This actually is a quartz device containing three elements which produce reflections that are dependant on the polarisation angle. As the returning beams have a different polarisation from the outgoing ones, they are reflected along a different path from this point on, and so do not return back through the grating to the laser.

Instead they travel through an astigmatic lens which produces a round spot when in focus on the photo-diode array. The reason for this lens we shall see later. Finally, the photo diodes convert the light and dark modulations into an electrical data stream.

THREE-BEAM TRACKING

The pickup assembly is moved over the underside of the disc by a servo motor. As beam tracking error must be within 0.1μm, the tracking must be extremely accurate and can only be controlled by the track itself. This is where those two auxiliary beams come into the picture.

They produce spots, one 20μm in advance and the other 20μm behind the main spot, on the silvered surface, but they are offset so that when the main spot is centered on the track one auxiliary is reading along the left hand edge of the track while the other reads along the right. (Fig.3).

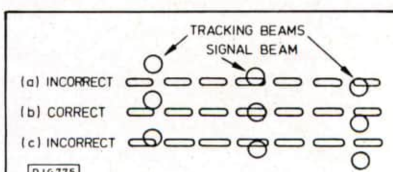

Fig.3. *Tracking beams should run either side of track (b). If incorrect one reflects a darker beam than the other so producing an error signal*

While the array is centered, the amount of modulation reflected from each auxiliary spot is equal. But if it strays to one side, one spot starts reading more of the track so reflecting a darker beam, while the other reads more of the space between the tracks, thereby returning a brighter beam.

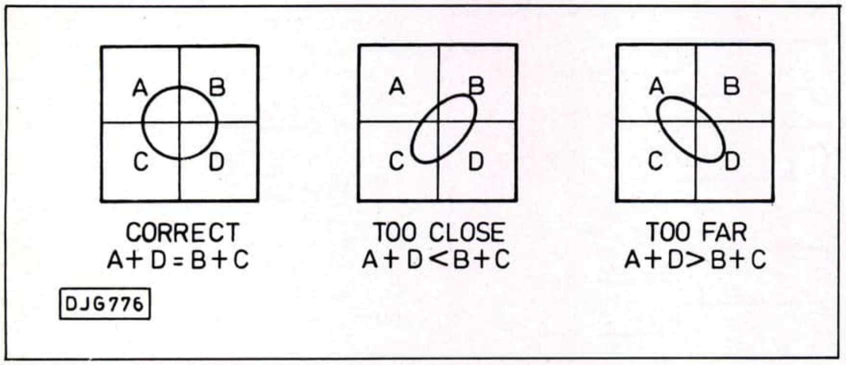

Fig.4. *Out of focus beam is rendered into an ellipse by the astigmatic lens. This illuminates diagonal pairs of photodiodes unequally. These are compared to produce a focus control signal*

The two beams are focused on their own respective photo diodes and the output from these are compared so that an error signal is produced when they are unequal. This is used to control the tracking motor. If either of the tracking beams are obscured temporarily by a disc surface blemish, the motor stops and the pickup remains stationary until both beams are again sensed. If this was not done, the circuit could generate a large spurious error signal due to the blacked out beam and so swing the pickup way off course.

Every track contains information as to its track number and timing from the start of the disc, so the decoder 'knows' what track is being read at any instant. This information can be displayed, or instructions can be given by the user to seek a particular track. This instruction puts the tracking motor into a fast mode which reverts to normal speed when the desired track number is detected. Just how that track and other information is recorded we will see in the next article.

FOCUSING

Focusing is another important function. Spot size is 1μm which is achieved with a lens aperture of 0.45 and a laser wavelength of 0.8μm. Defocusing enlarges the spot which could thereby read more than one hump at the same time causing corruption of the data signal.

The depth of focus of a 0.45 lens is 4μm, which makes focusing very critical, but here it is aided by the beam passing through the material of the disc. The material has a refraction index of 1.5, which is the ratio of the speed of light through a material compared to that through a vacuum. Resulting refraction increases the depth of focus. An advantage of a small depth of focus is that as the disc surface is 1,200μm from the silvered layer, it is a long way out of focus and so surface blemishes cast only a diffused image. They are thus less likely to cause errors.

Focusing error must be within 0.5μm, yet disc warp may cause a displacement of the surface of up to 500μm. Focus must therefore be continuously and automatically variable. The lens system which consists of four elements has a concentric coil surrounding it, and it is free to move vertically between the pole pieces of a permanent magnet. Movement is caused by current passing through the coil in the same manner as a loudspeaker cone.

There are four signal photo diodes arranged in a square block configuration. The spot reflected from the disc is focused at the centre of the four and when the system is in focus the spot is round so illuminating equal portions of all the diodes. If it goes out of focus, the spot develops into an ellipse due to the presence of the astigmatic lens, and thereby illuminates a greater area of one diagonal pair than the other.

When the disc surface is too close, the ellipse lies across the second and third photo diode (B and C in Fig. 4) but when it is too far it straddles the first and fourth (A and D). So, the four are connected so that the output of each diagonal pair is added and compared to that of the other diagonal. (Fig. 5).

The comparator thereby produces an error signal which is amplified and fed to the lens coil. The polarity of the error signal, hence the direction of the lens movement, depends on which diagonal pair has the greater illumination. All four are summed to produce the main data stream signal.

SINGLE-BEAM SYSTEMS

If the two tracking beams are dispensed with the optical system can be simplified. The optical grating, Wollaston prism and quarter wavelength plate can be eliminated. The complete optical system can be contained in a simple, easily replaceable plastic-cased unit which can be swung across the disc on a pivot like a pick-up arm. A linear carriage is needed for three beams to maintain their tangential alignment. Furthermore, the whole energy of the laser is applied to the signal beam instead of 50% as with three beams. There are two methods of focusing and tracking. (Fig. 6).

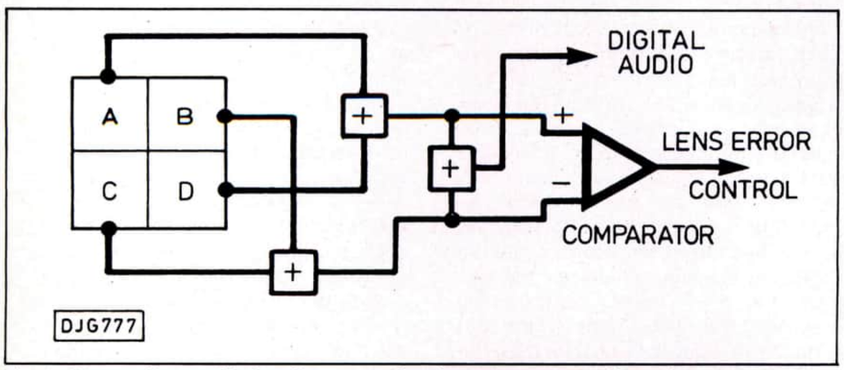

Fig.5. *Diagonals are added and compared with opposite pair to derive focus correction signal. All four photodiodes are summed for main audio data signal*

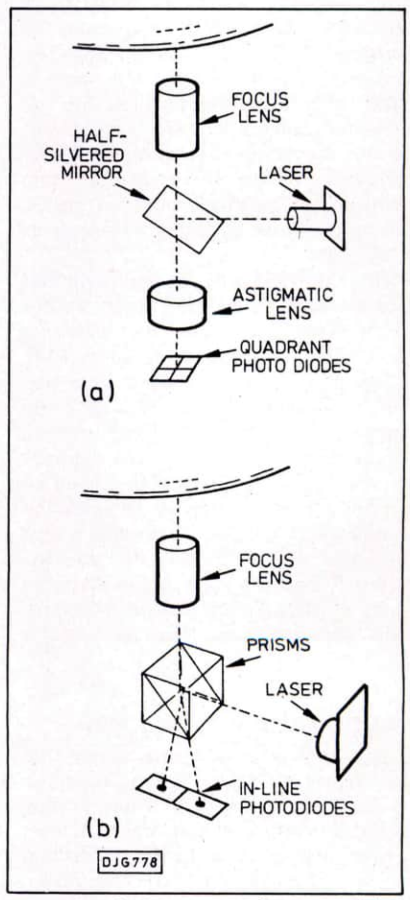

Fig.6. *Single beam optical systems (a) Quadrant diodes (b) In-line diodes*

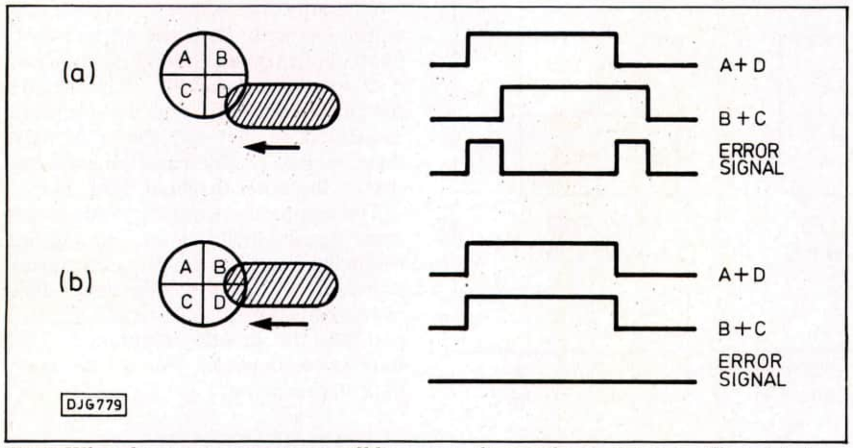

Fig.7. *When hump encounters an off-line spot, diagonal pair A+D is affected before B+C. Voltage at the comparator is thus in advance of B+C and a pulse is produced which is used as an error signal (a). When the spot is on-line (b), both diagonal pairs are affected simultaneously and no pulse appears*

Optical Path. The laser beam is reflected by an angled half-silvered mirror through the focus lens. After reflection from the disc it re-enters the lens, passing through the mirror and final lens to the photo diodes.

Quadrant Diode System. With this the focusing is the same as with the three-beam system, but the four quadrants of the detector are also used to detect deviation of the spot from the centre of the track.

This is how it is done. If the beam is off-course, one half of the spot reads the humps while the other reads the land so that one half of the reflected beam is darker than the other, and likewise the spot produced on the photo diodes. But as the quadrants are connected diagonally, both have one quadrant reading humps while the other reads land. Hence the output from both pair of diagonals is the same.

There is a difference though, and that is in the timing. (Fig. 7). A hump reaches the top quadrant of one pair before it encounters the bottom quadrant of the opposite one. So the output from one pair of quadrants arrives at the comparator just before that of the other, resulting in a pulse. Its polarity depends on which pair is affected first, and so indicates to which side of the track the spot has deviated. An error signal is thus produced which controls the tracking motor. When the spot is on line, the hump reaches the top quadrants of both pairs simultaneously, and no pulse is generated.

IN-LINE DIODE TRACKING

The difference with this system and the quadrant one is that the reflected beam from the disc is split into two by the final prisms in the optical path. Two prisms are combined, the exit one having a wedge shape which divides the beam. Their joint surface is half-silvered by an evaporated film that serves to reflect the incoming laser beam up through the focus lens. Although split into two, the system is termed single beam because only one beam is applied to the disc.

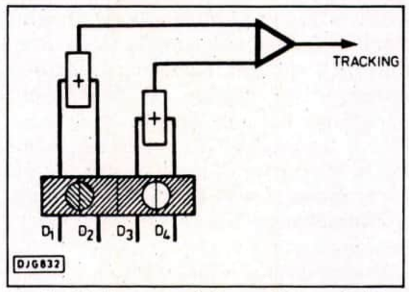

Fig.8. *Light spots fall equally on in-line diodes D1 and D2, D3 and D4. When off-centre one spot becomes darker than the other. Outputs of each pair are added and compared to provide tracking error signal*

Both beam halves fall on four photo diodes which are arranged in a row. (Fig. 8.) One falls equally between D_1 and D_2, and the other between D_3 and D_4. If the spot on the disc is off-centre, the side of the beam reflected from the adjacent land is brighter than that reflected from the humps. So the two beam halves differ in brightness and there is a difference in output level between the diodes $D_1 + D_2$, and $D_3 + D_4$. This is used as an error correction signal.

However, dirt on the lens or prism, or other defects could produce permanent brightness inequality between the two beam halves resulting in tracking error. To avoid this a second tracking error signal is generated by a applying a 600Hz current to the tracking coil. This makes the arm oscillate from side-to-side, displacing the beam by $\pm$ 0.05μm. As any deviation from true centre varies the respective brightness between the two beam halves, both are thereby modulated by a 600Hz signal. If the spot drifts to one side, the signal increases in one pair of diodes and decreases in the other.

Summing the signal from both pairs produces either a positive or negative control signal depending on which side of the track the spot has strayed. This serves to 'fine tune' the first and main control signal.

FOCUS

When the spot is sharply focused on the disc, two sharp images appear on the photodiodes, one in between D_1 and D_2, and the other between D_3 and D_4. If the spot goes out of focus, the images also become diffuse and move closer together or further apart depending on the direction of the focus error. Fig.9. Thus the inner pair of diodes D_2 and D_3, have a different illumination level than the outer pair D_1 and D_4. A comparison provides a difference error signal that is applied to the focus coil.

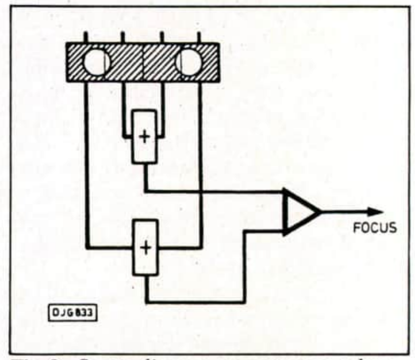

Fig.9. *Spots diverge or converge when beam is out of focus. Inner and outer pairs are added then compared to provide the focus error signal*

MOTOR SPEED CONTROL

The disc motor runs more slowly when scanning the outer tracks, than when the inner ones are read. Thus the data comes off the disc at constant speed. But unlike gram motors which need to be speed controlled to very fine limits to avoid wow and other pitch fluctuations, the cd motor speed is not particularly critical.

After emerging from the decoders the data is loaded into one end of a memory. It is clocked out at the other at constant speed by a quartz clock generator and fed to the d/a converters. It doesn't matter much if the input to the memory varies in speed as long as it is not too fast and so fills the memory completely to overflowing, or too slow so that it empties it.

The situation is rather like making regular weekly withdrawals from your bank account, the input may be irregular but as long as there is always sufficient to keep you in the black it doesn't matter very much. In that case of course it is not a bad thing to put so much in that the balance piles up!

With the cd output memory it is kept about 50% full so that there is room in both directions for adequate regulation. The amount it contains is continually measured and if it begins to fill up beyond the 50% level, this indicates that the motor is running too fast. Or if the level drops, the motor is too slow. A correction signal is thereby derived and used to change the motor speed.

So minor speed fluctuations are of no consequence, as the output frequency is firmly controlled by the quartz clock. Wow is therefore non-existent with a cd player.

We have come to the point where the light modulations produce corresponding electrical signals from the photodiodes. It may be thought that these are just digital signals that correspond to the original audio, and that all that is now required is to feed them to a d/a converter to recover that audio.

There is much more to it than that. The signal must be converted so that it can be read by a spot much larger than the smallest digital unit, (like using a gram stylus several times larger than the record groove). Correction elements must be included so that if parts of the signal are missing or incorrect there is no audible effect. Track numbering and other data needs to be included, and the two stereo channels must be combined in one track in a manner so that they can be parted with a much larger separation than is possible with an lp record.

How is it all done? Watch out for next month's issue and find out! **PE**