Phys. Educ. **28** (1993). Printed in the UK

The physics of the compact disc

John A Cope

The compact disc provides an excellent example of the application of basic physics principles, from the derivation of the signal by interference to the dependence of audio frequency response and playing time on the optical wavelength used.

Compact discs have been on sale to the consumer for ten years, and during that time they have taken over the major part of the recordings market. Their success has been partly due to the greater quality of sound that they make possible. It has been claimed (Dinsdale 1992) that LPs give better reproduction but only on equipment costing many times that of the CD player. My Compact Disc player enables me to enjoy music without the irritation of Tape Hiss or Record Crackle, which could so easily spoil the quieter passages, and the wide dynamic range and the absence of wow in the reproduction are further bonuses. As a music lover I enjoy listening to my Compact Discs and, as a physicist, I appreciate the significant contribution made by physics to the concepts and design of the components of the playing system.

When Compact Disc players were first produced they were at the leading edge of audio technology, involving the combination of the skills from a number of disciplines. Laser technology, geometrical and physical optics, computer science, electronic and mechanical engineering all contributed to their development. Not only was there a marriage of all these skills but the production at that time required a unique collaboration between Sony and Philips.

Physics is the fundamental discipline behind many of those skills listed above, and this article will describe the application of physics principles to the dimensions and operation of the components of the Compact Disc player—examples that are useful in the teaching of physics.

This article follows the light beam through the system from the laser light source, via the disc to the photodiode detectors.

Basic principles

The music is recorded on the disc as steps arranged along a track spiralling out from the centre of the disc. In the presence of a step, light reflected from its top will destructively interfere with light reflected from the area surrounding the step. In the absence of a step no such destructive interference occurs. Thus the intensity of the received signal varies as the steps pass under the laser beam. The lengths of the steps and of the gaps between them transmit a binary type code, which is processed to reconstitute the original sound.

The light source

Since the system works on the interference of light reflected from the playing surface it is essential to have a coherent source of monochromatic light. The light source must be small and reasonably robust. A semiconductor laser is the obvious device to employ; an aluminium gallium arsenide laser giving light of wavelength 780 nm is commonly used. One problem with semiconductor lasers is that their output is sensitive to temperature changes. Variations in output are compensated by a negative feedback process in which a photodiode detects the light output from the laser and adjusts the current to the semiconductor laser appropriately.

The recent development of a blue light laser (Haase *et al* 1991) has implications for the future design of compact and optical disc systems. The shorter wavelength that would be used to read the data from the disc would allow narrower and more

---

John Cope is a Clinical Scientist with Clwyd Health Authority and a Tutorial Fellow at the University of Wales–Bangor. He gained his doctorate from St Andrews in 1967, and for 20 years was Director of Studies in the Science Department of North East Wales Institute.

0031-9120/93/010015+07 $07.50

densely packed recorded tracks, allowing longer playing times.

The focusing lens

In order for the disc to contain enough information for a reasonably long playing time the data must be compressed onto very narrow tracks on the disc. This then requires the beam, used to read the information, to be focused to a sufficiently narrow spot so that it does not overlap adjacent tracks. This imposes stringent requirements on the lens focusing system.

Unlike the LP the music track on a CD is 1.2 mm beneath the surface of the disc. A short focal length lens is used to converge the laser light through the surface of the disc towards the information layer (figure *1*). This has the advantage that the beam will be relatively broad as it passes through the surface of the disc. Scratches or marks on the surface of the disc will be out of focus and, if they are not too wide, will allow enough light to penetrate to the recorded track for the modulation in the reflected light to be decoded.

For a typical Compact Disc player the lens produces a cone of light converging onto the information layer with a semi-angle U at its vertex of about 27°. The material of the disc (often a polycarbonate plastic) has a typical refractive index of 1.55. Refraction as the light enters the surface of the compact disc will reduce the angle U' of the cone of light inside the plastic to 17°.

The information layer lies 1.2 mm beneath the surface of the disc and therefore the diameter of the disc of light entering the surface will be $2 \times 1.2\,\text{mm} \times \tan 17° = 0.73\,\text{mm}$. This is several orders of magnitude higher than the width of the spot required to scan the recording track. A scratch or a hair on the surface, or a line drawn with a 0.5 mm nib pen would not completely block the passage of the beam to the music track and would not then affect the music output.

Diffraction effects

Even if the lens has no aberrations, diffraction effects will still prevent focusing of the beam to a very small spot. The distribution of light intensity about the central maximum due to diffraction through the circular aperture of the lens produces an Airy ring diffraction pattern (Longhurst 1957). This pattern contains a bright central disc carrying approximately 84% of the total energy, surrounded by rings of rapidly decreasing intensity (figure 2). The radius R of the central Airy disc produced by light of wavelength λ in a medium of refractive index n is given by the equation

$$R = \frac{0.61\,\lambda}{\sin U} = \frac{0.61\,\lambda}{n \sin U'}$$

where, as above, U is the semi-angle of the converging cone of light in air, typically 27°, and U' is the semi-angle inside the disc material of refractive index 1.55 (= 17°). If λ is the wavelength of the light in a vacuum, λ/n is the wavelength in the medium of the coating of the disc. The wavelength of 780 nm in air becomes 503 nm in the plastic material of the disc. The radius of the Airy disc in the plastic for the 17° semi-angle of the converging cone of light is 1.05 µm. It is this dimension that dictates the minimum separation of the recorded tracks. Adjacent tracks are 1.6 µm apart, which

Figure 1. The strongly converging beam is broad as it enters the surface of the disc.

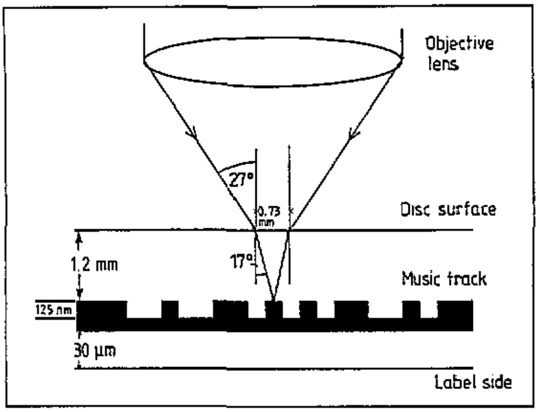
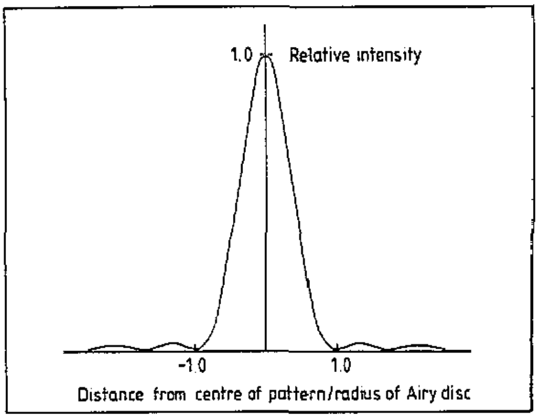


Figure 2. The distribution of light due to diffraction through a circular aperture.



puts one track beyond the minimum in the Airy ring pattern of the beam scanning the neighbouring track. Track separation is made greater than the theoretical minimum to allow for the broadening of the spot due to lens aberrations and to allow for some wandering of the spot as it follows the length of the track (tracking error). The diameter of the spot of light incident on the recorded track will also be affected by the distance of the track from the lens. Ideally this equals the focal length of the lens, but if the disc is not perfectly flat or if it does not sit perfectly in the player the distance could vary. The depth of focus is the range of distances over which the spot of light is acceptably narrow. Variations in focal spot size affect the proportion of light reflected from the steps relative to that reflected from the background. In an extreme case the spot could become so large at the playing surface that it would overlap two adjacent tracks. The lens–disc surface is maintained constant to ±2 μm by a servo mechanism.

The steps on the recording track used to carry the music code are 0.5 μm wide. The length and spacing along the track of the steps is used to encode the signal, with a minimum step length of 0.833 μm. If the 1.05 μm radius spot of light is centred on a long step of width 0.5 μm the step would reflect approximately 50% of the incident light. The reflections from the area around the step will therefore have about the same intensity as the reflections from the step itself, and the destructive interference between these two reflections will have the maximum effect on the modulation of the resultant reflected beam. In practice this is rarely achieved because of lens aberrations and because the beam would not be perfectly centred on the track. Nevertheless the modulation is still adequate for transmitting the binary code to the detector.

An electron micrograph of the surface of a CD is included in the preceding article by M Cornwall.

The information layer

Destructive interference occurs if the path difference between a ray reflected from a step and one reflected from the surrounding background surface is half a wavelength. The wavelength of the light in the medium is 503 nm and therefore a path difference of about 250 nm is required. This path difference is achieved in the outward and return journey of the two rays by a step of height 125 nm (i.e. a quarter wavelength). In the preparation of the disc a glass former is coated with photoresist to a depth between 110 nm and 135 nm. These variations from the ideal quarter wavelength will affect the step depth and therefore the completeness of the destructive interference process, and hence the modulation in the reflected signal. However, the modulations are only used to transfer a digital code to the receiver and it is only the temporal pattern of peaks in the signal—not their amplitude—that is important. Therefore small variations in step height are acceptable. A laser beam of wavelength 350 nm is used to expose the photoresist layer, which when developed gives pits of the required depth. There are many stages before the consumer's disc is pressed but the overall result is the production of a 'mate' to the master disc having steps corresponding to troughs in the original master. The plastic disc has the pattern of steps pressed into it and the whole surface is then coated with a thin layer of aluminium to make it highly reflective. A thin lacquer coat is then applied for protection, onto which is printed the label. The disc is read from the face opposite to the label.

Playing time

The information track spirals out from the centre of the disc and, apart from the lead-in and lead-out tracks, the disc playing surface extends from an inner radius of 25 mm to an outer radius of 58 mm, giving a programme area 33 mm wide. The number of tracks, N, is given by

$$N=\frac{\text{width of programme band}}{\text{distance between tracks}}=\frac{33\times 10^{-3}\,\text{m}}{1.6\times 10^{-6}\,\text{m}}$$

$$=20\,625.$$

The total length of track on a typical CD, L, is given by

$$L=\text{number of tracks}\times\text{average circumference}$$

$$=20\,625\times 2\pi\times\frac{(25+58)\times 10^{-3}}{2}=5.38\text{ km}.$$

The disc rotates so that the readout spot moves over the track at a speed of 1.2 m s^{-1}. This results in a maximum playing time of

$$\frac{5.38\times 10^{3}\,\text{m}}{1.2\,\text{m}\,\text{s}^{-1}}=4483\,\text{s}=75\text{ minutes}.$$

A longer playing time could be achieved by using a slower tracking speed but this would limit the frequency response of the system. The length of the steps is determined by the coding system used. The details are too lengthy to go into here but the 8-bit information obtained from digitizing the audio signal is not directly coded onto the disc as a series of 1s (steps) and 0s (no step) because this could lead to a very large number of closely spaced steps with the attendant risk of confusion between similar code patterns. The 8-bit code is transformed into a 14-bit word (EFM—eight to fourteen modulation) (Watkinson 1985) using a look-up table which only allows patterns with a minimum of 3 consecutive zeros and a maximum of 11 zeros. These sequences of zeros are then converted into steps on the track. So the shortest step corresponding to 3 zeros has a period equal to 3 × (period of the master clock). The master clock frequency is 4.3218 MHz, giving a period of 231.4 ns, and a '3-zeros' step at a scanning speed of $1.2\ \mathrm{m\ s^{-1}}$ would give a step length of

$$1.2\ \mathrm{m\ s^{-1}} \times 3 \times 231.4 \times 10^{-9}\ \mathrm{s} = 0.833\ \mu\mathrm{m}.$$

Other information must also be coded into the track for the purposes of synchronization and error correction. The audio data words and the synchronization/error correction words are combined in frames consisting of 588 bits. Of these 588 bits only 192 bits carry the direct audio information. So the audio bits are read at a frequency of

$$\frac{192}{588} \times 4.3218\ \mathrm{MHz} = 1.4112\ \mathrm{MHz}.$$

These bits are combined into 16-bit bytes, which means the sampling frequency is

$$\frac{1.4112}{16}\ \mathrm{MHz} = 88.2\ \mathrm{kHz}.$$

This gives a sampling frequency of 44.1 kHz per stereo channel. The Nyquist theorem (Pohlmann 1989, see also Baert *et al* 1988) states that the sampling frequency must be at least twice the maximum sound frequency to be recorded. A lower sampling frequency would lead to aliasing, i.e. the production of false low frequency components in the output. According to this theory a system using a sampling frequency of 44.1 kHz should be able to reproduce faithfully frequencies up to 22 kHz. This is more than adequate for recording sound since the upper threshold frequency for human hearing is about 20 kHz.

We still need to show that the optical readout system is capable of dealing with these data acquisition rates. Because of the eight-to-fourteen modulation coding algorithm the shortest step-gap cycle is 3 + 3 clock periods long = 6 × 231.4 ns = 1.4 μs, which corresponds to a frequency of 720 kHz. The Airy disc has a radius of 1.05 μm and moves at $1.2\ \mathrm{m\ s^{-1}}$ relative to the track. Two peaks would be just distinguishable if the maximum of one coincided with the first minimum of the other. This leads to a cut-off frequency of $1.2\ \mathrm{m\ s^{-1}}/(1.05 \times 10^{-6}\ \mathrm{m}) = 1.14\ \mathrm{MHz}$. At frequencies above this threshold successive steps on the surface would not be distinguished, i.e. the modulation of the output would be zero. Below the threshold frequency the output rises approximately linearly. The maximum frequency encountered is 720 kHz, which allows the system to cope adequately.

The pickup

The beam reflected from the information layer must be deflected towards the detecting system without obstructing the incident beam. This reflected beam must not only carry the audio information but is used to detect and then control the tracking and focus of the beam on the data track. The design of the pickup varies from one manufacturer to another. One pickup design called a three-spot system, involving the greatest number of physics principles, will be described here (figure 3).

Light from the semiconductor laser first passes through a diffraction grating. The zeroth-order beam emerging from the grating is the one that is used to read the information from the recorded track. The first-order beams on either side of the main beam are used to detect any wandering from the line of the track. The three rays of the incident beam then pass through a polarizing prism, making the beam plane polarized. This plane-polarized beam then passes through a quarter wave plate, i.e. one that introduces a 90° phase difference between the two components of the electric field vector. The beam is now circularly polarized and remains so until it returns after reflection from the disc. Passing through the quarter wave plate a second time introduces a further 90° phase difference between the components, which makes the beam polarized in a plane 90° rotated from that of the beam emerging from the polarizing prism (figure 4). The plane of polarization is now such that the polarizing prism efficiently reflects the returning beam towards the receiver photodiodes rather than transmitting the beam back towards the laser. A cylindrical lens introduces some deliberate astigmatism into the system before the central spot falls onto a four-quadrant photodiode (figure 5). If the disc is correctly positioned relative to the objective lens the central spot on the photodiode array is circular and all of the quad-

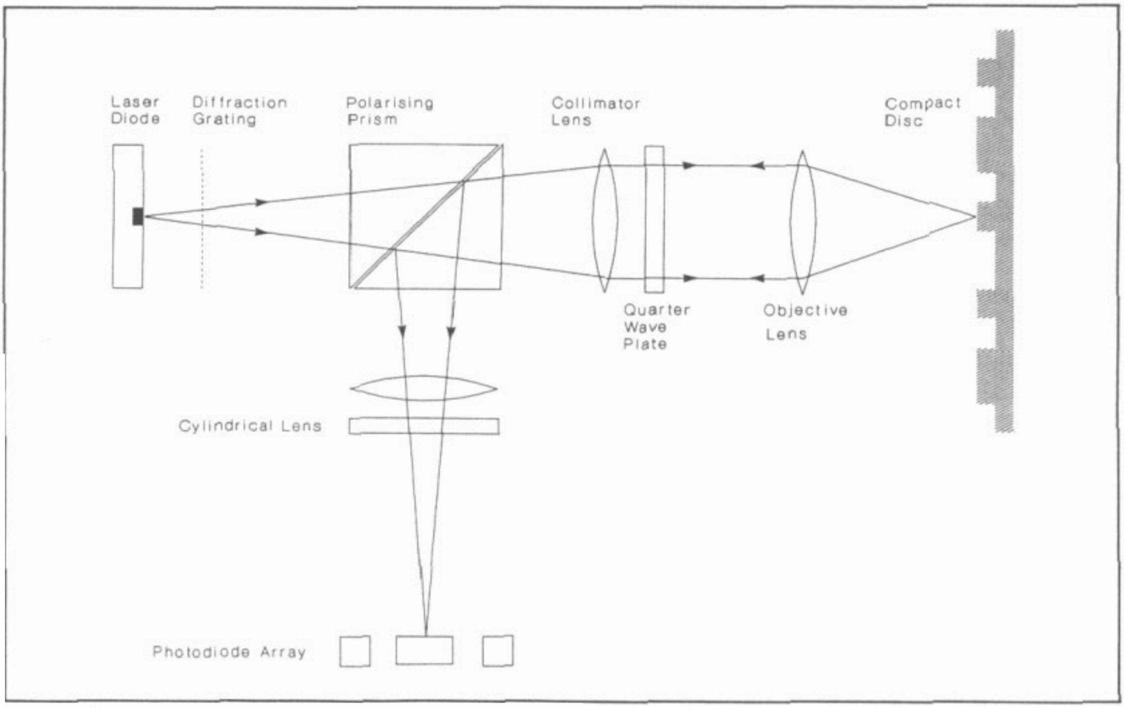
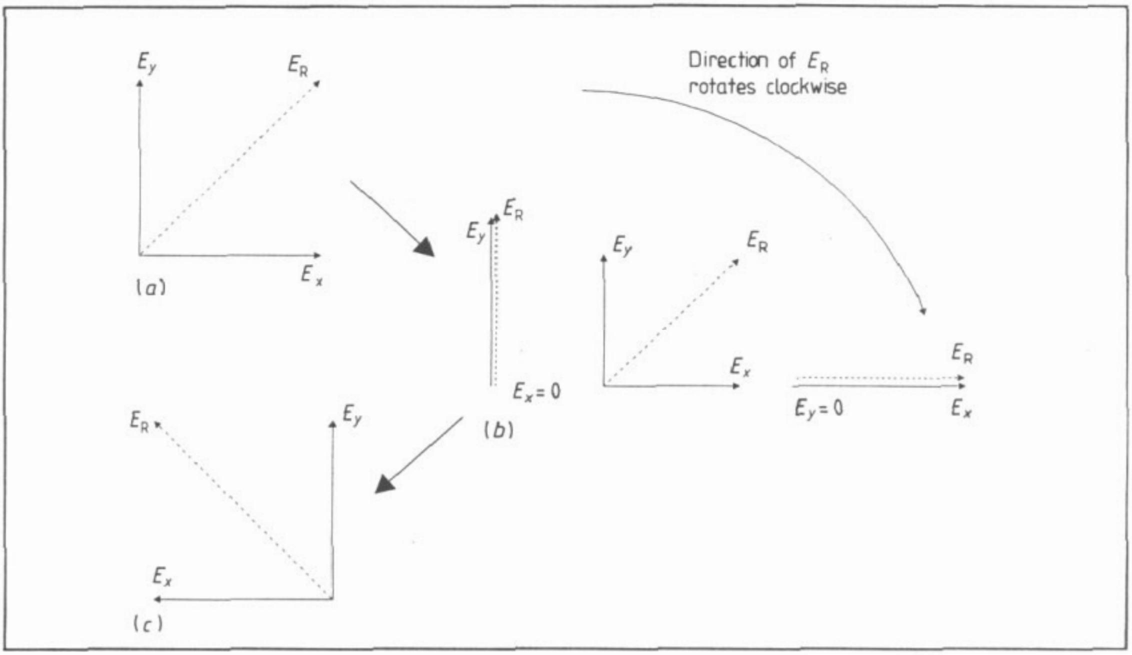


Figure 3. Schematic diagram of the main components of the system used to read the data from the disc.

Figure 4. The polarization changes involved in the three-spot pickup system. (*a*) E_x and E_y in phase, plane polarized light. (*b*) E_x and E_y 90° out of phase, circularly polarized light. (*c*) E_x and E_y 180° out of phase, plane polarized light with plane shifted 90° from that in (*a*).



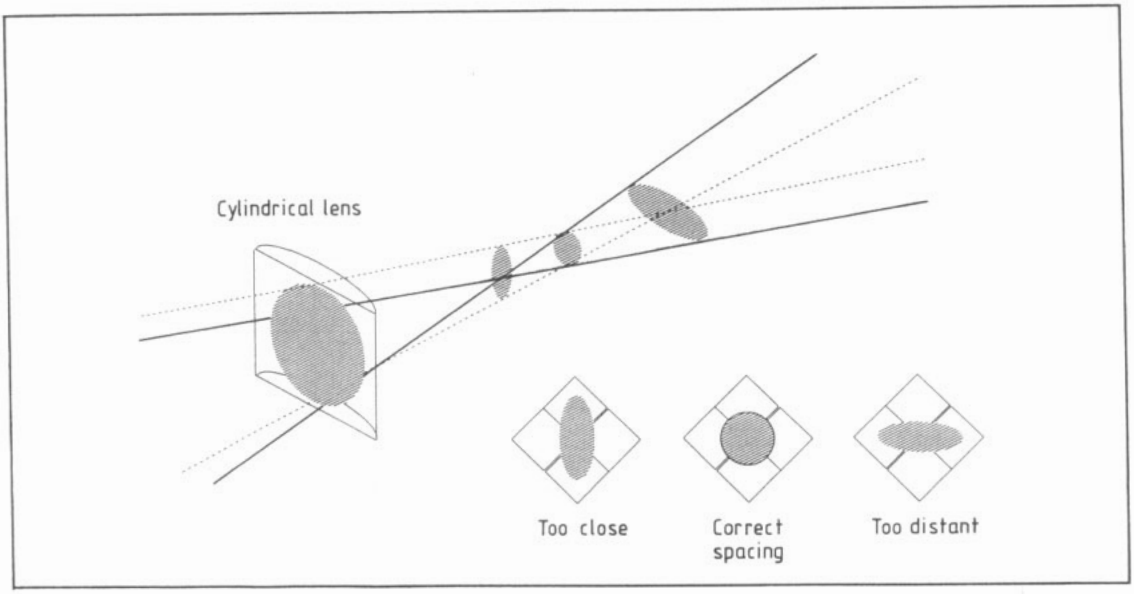


Figure 5. A cylindrical lens introduces astigmatism into the optics. The shape of the cross section of the beam changes on moving through the principal focus.

rants receive equal signals. If the disc–lens distance changes due to warping or other mechanical problems the astigmatism in the system will cause an elliptical spot to fall onto the quadrant. The signal from one diagonal pair will be stronger than from the other pair. The difference signal is sent to a servo mechanism to adjust the focus appropriately. The lens adjustment mechanism is similar to the structure of a loudspeaker, which allows the lens position to be adjusted rapidly (figure 6).

The two side spots of the diffraction pattern are brought to focus on two separate detectors flanking the central quadrant photodiode. If the spot B (figure 7) reading the music track is accurately centred on that track, the two side spots A and C will fall on the flat area of the disc either side of the track. In this situation the central spot will show greater modulation than the light to the two side spots. If, however, the lens mechanism drifts to one side of the track the modulation of the three beams will change as shown in the diagram. An increase in the modulation of the light in either side spot indicates a drift from the true track, and the signal from the three detectors can then be fed to a servo mechanism to correct for that drift.

Thus the signals from the six photodiodes are combined in different ways to provide audio, focus correction and tracking correction signals (figure 8).

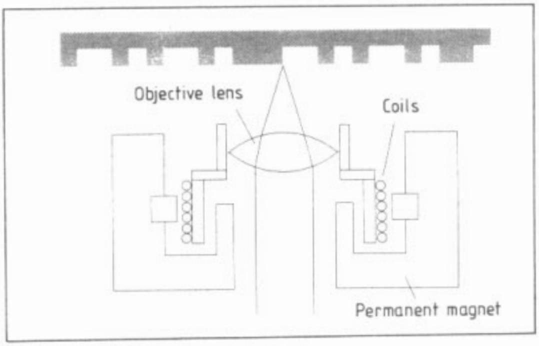


Figure 6. The objective lens is mounted similarly to the core of a loudspeaker, allowing rapid position corrections.

Conclusion

When you are relaxing to the music from your Compact Disc system give some thought to the physics principles involved in the operation of its component parts: the principles of interference used to achieve the modulation in the received beam; both the geometrical optics of the lens system and the physical optics associated with diffraction through the lens aperture which determine the width of the beam at the recorded track and hence the total playing time possible; the use

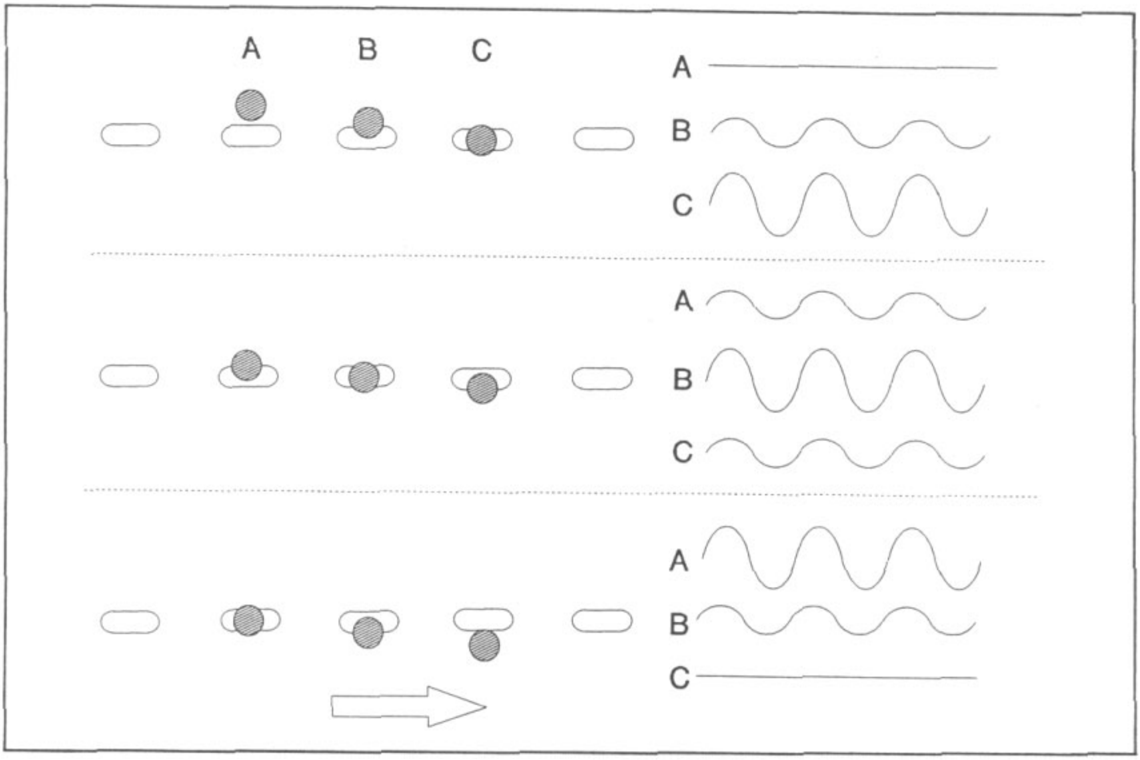


Figure 7. Top and bottom: the central spot B has drifted to one side of the music track and the modulation is greatest in one of the side beams A or C. Centre: The central spot B is correctly located over the track and the modulation from the central spot is a maximum.

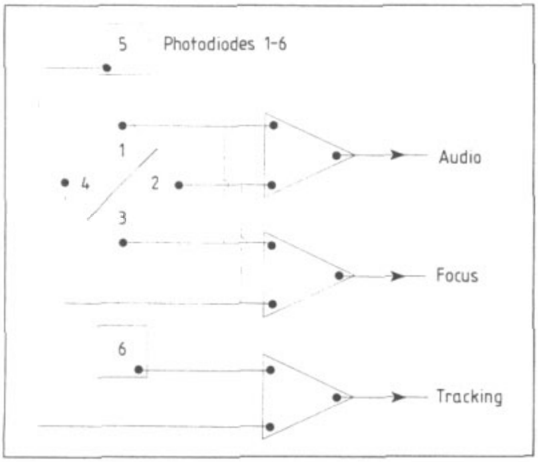


Figure 8. Combining signals from the six photodiodes in different ways provides the output for the audio, focusing and tracking circuits.

of a quarter wave plate and the rotation in the plane of polarization used to separate the incident beam from the modulated beam reflected from the recorded track.

CDs have been a phenomenal success. The sale of CDs has for some time outstripped that of LPs. The CD provides an opportunity to discuss a large number of physics principles concerned with a piece of equipment involving modern technologies that young people of all tastes can appreciate.

References

Baert L, Theunissen L and Vergult G (ed) 1988 *Digital Audio and Compact Disc Technology* (London: Heinemann Newnes)

Dinsdale J 1992 How hi was my fi? *Phys. World* **5** August

Haase M A, Qiu J, DePuydt J M and Cheng H 1991 Blue-green laser diode *Appl. Phys. Lett.* **59** 1272–4

Longhurst R S 1957 *Geometrical and Physical Optics* (London: Longmans)

Pohlmann K C 1989 *The Compact Disc, A Handbook of Theory and Use* (Oxford: Oxford University Press)

Watkinson J R 1985 Channel code and disc format *Electronics and Wireless World* May