PHILIPS

Service
Service
Service

Basic course Compact Disc

COMPACT
disc
DIGITAL AUDIO

Instructor's manual

For internal use only

Concern Service Centre
Supporting Services
Training Group

References

Electronic Components and Applications

Vol. 2, No. 4, August 1980
Monolithic 14-bit DAC with 85 dB S/N ratio
R.J. v.d. Plassche

Vol. 4, No. 3, May 1982
ICs for Compact Disc decoders
J. Matull

All enquiries should be addressed to:
Nederlandse Philips Bedrijven B.V.
Concern Service
5600 MD Eindhoven
The Netherlands

Printed in The Netherlands

Preface

This manual has been developed by Concern Service in close cooperation with Audio Service and the Audio Division's Compact Disc Development Laboratory.

The contents of this manual provide basic and background knowledge for service instructors training technicians on the Compact Disc.

Each chapter is preceded by a number of questions which refer to the subjects treated in the chapter concerned.
If you are able to answer these questions, it will suffice to only go over to the summary.

Contents

Introduction

The progressive development of technology has, in its turn, made it possible to bring the storage and reproduction of sound to greater perfection.
At present it is possible to obtain good sound reproduction from high-quality discs.
However a number of difficult problems still remain in the conventional system of mechanical scanning. These are:

- disc wear
- damage to the discs
- signal-to-noise ratio
- dynamic range
- channel separation
- wow and flutter

The Compact Disc system offers a solution to these problems.
Since the disc is read out optically there is no mechanical contact between the pick-up unit and the disc.
Thus, the occurence of wear is prevented.

The information on the disc is covered by a relatively thick layer of transparent plastic so that scratches, dust and dirt cannot damage it.
Furthermore, the laser beam that scans the disc is focused relatively far below the plastic surface so that dust, scratches and dirt will hardly (if at all) affect the ultimate audio signal.

Digitization of signal processing has resulted in a signal-to-noise ratio and a dynamic range better than 90 dB for the Compact Disc system.
Digitization has also made it possibile to correct errors which can occur during signal transmission. For this purpose, the digital information on the disc is supplemented by additional codes. These codes enable the player to check and, where possible, correct the digital information.

Audio information of one channel cannot affect the information contained in the other channel when applying the Compact Disc system, because the signals are processed completely separately.

This ensures an impressively good channel separation.

Wow and flutter belong to the past; within defined limits, processing of the signals occurs independently of variations in the turntable's rotational speed, the reference being a crystal oscillator.

All these good qualities make the Compact Disc a unique device with an extremely good quality of reproduction.

General block diagram

Questions

1. **What is an optical pick-up unit composed of?**

How is an analogue signal converted to a digital signal?

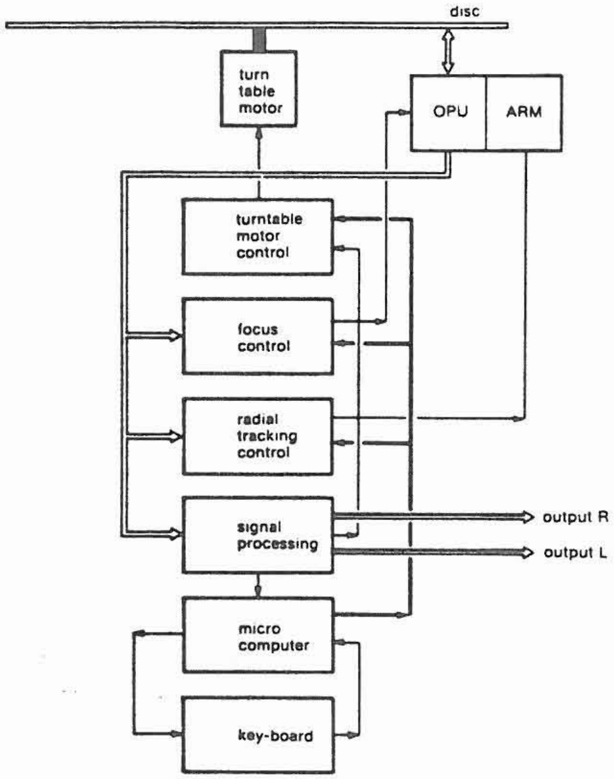
What must be added to the digital signal to enable the correction of errors?

4. **How is the string of 0s and 1s recorded on the disc?**
5. **Wat are the functions of**
   - **DEMOD?**
   - **ERCO?**
   - **CIM?**
   - **FIL?**
   - **DAC?**
6. **How is the radial error signal obtained?**
7. **How is the rotational speed of the motor controlled?**
8. **What provision has been made to keep the laser beam focussed on the reflective surface of the disc?**

Description of the general block diagram

Introduction

A distinctive feature of the Compact Disc system is that the information contained in a reflective disc is read out by means of an optical system.



The information is present in digital form, and represented by two electrical levels only, i.e. 0 volts and +5 volts.
Inside the player the information is demodulated and existing errors are, whenever possible, corrected before a digital-to-analogue converter reconverts the digital signal into a reproduction of the original analogue audio signal.

The optical system is composed of a laser, a set of lenses and light-sentitive diodes (photodiodes).
The lens system ensures correct focusing of the laser light on the disc surface, whereas the photodiodes convert the reflected light into electric signals.

The information is present on the disc in the form of pits or depressions, pressed into the reflective surface.
Light that hits these pits is scattered in such a way that the intensity of the reflected light falling on the photodiodes is much lower than that of the light reflected back from the reflective surface between the pits.
This results in variations of the electric signal that leaves the photodiodes.

During disc read-out the pick-up unit, comprising a laser and a set of lenses, will have to follow the track with great accuracy.
There exists, however, no mechanical contact between the pick-up unit and the disc to achieve correct mutual positioning.
By using more than one photodiode it becomes possible to measure the deviation of and hence correct the position of the arm holding the pick-up unit.

The speed at which the turntable is rotating also needs regular adjustment to ensure a constant rate of recovery of information. To achieve this, the speed at which the information is read from the disc is measured; if the rate is too high, the turntable is braked; too low a rate of information flow is corrected by increasing the turntable's rotational speed.

An essential requirement for the system's correct operation is that the laser light should be perfectly focused on the information track. The accuracy of focusing can be determined by means of the reflected light. If the system is out of focus, correction is achieved by controlling the objective. For this purpose, the objective is movable in a vertical direction and can be controlled electrically.

A microcomputer controls the operation of the player and also reads the keyboard, thus integrating all the functions.

The pick-up unit

The pick-up unit is composed of:

- a laser
- a prism
- a collimator lens
- an objective
- a few photodiodes

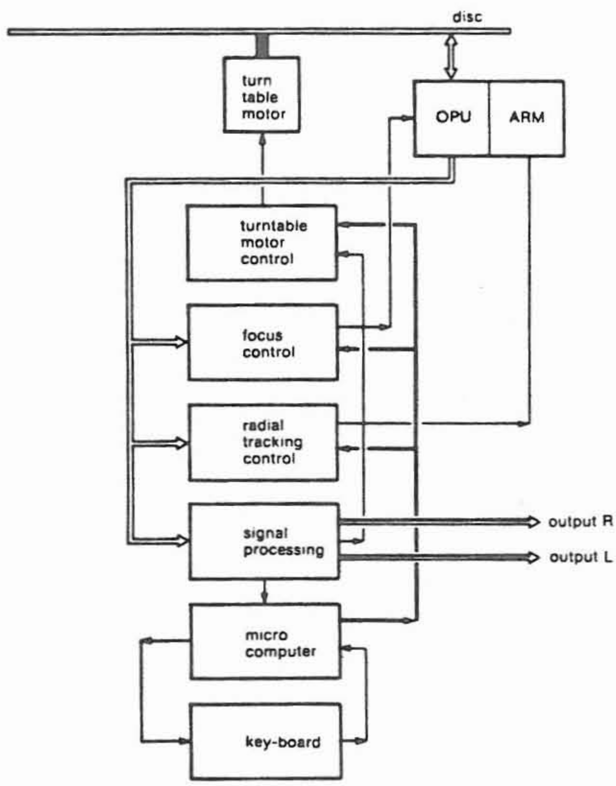
The laser generates a light beam which - via the other optical elements - finds its path to the disc, is reflected back and thus becomes the signal carrier.

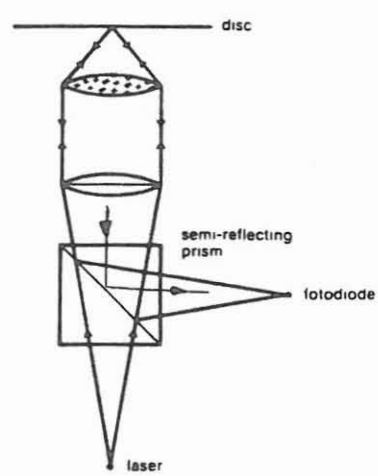
The prism offers undisturbed passage to the laser light going to the disc; the light reflected back from the disc surface is, however, deflected by the prism towards the photodiodes.

The collimator lens causes the laser beam to become a parallel light beam.

The objective's function is to focus the light beam on the disc.
The objective can be displaced in a vertical direction. If the distance between disc and objective is varying, the displacement facility makes it possible to maintain correct focusing by adjusting the objective's position.

The photodiodes convert the intensity-modulated light into an electric signal.
The player contains a total of 4 photodiodes. By suitably combining the signals from the photodiodes it is possible to obtain the audio information and to obtain also errorr signals to correct focus and radial tracking.

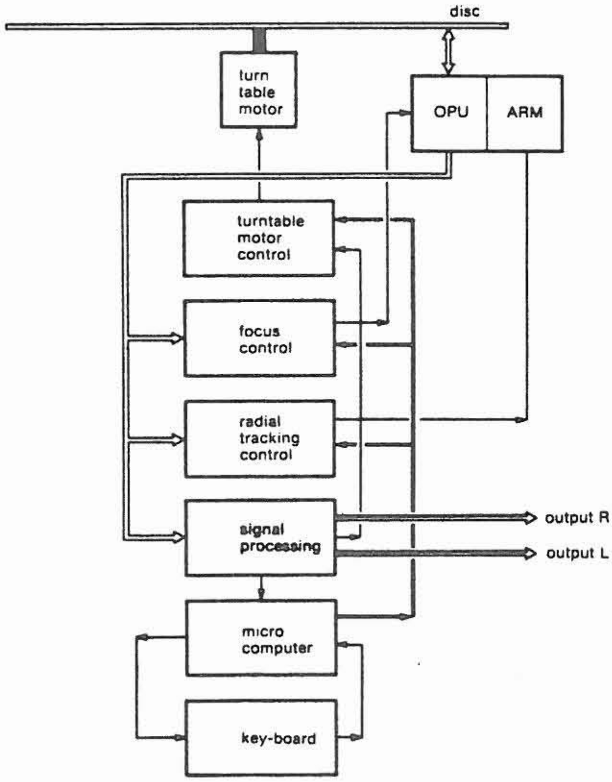




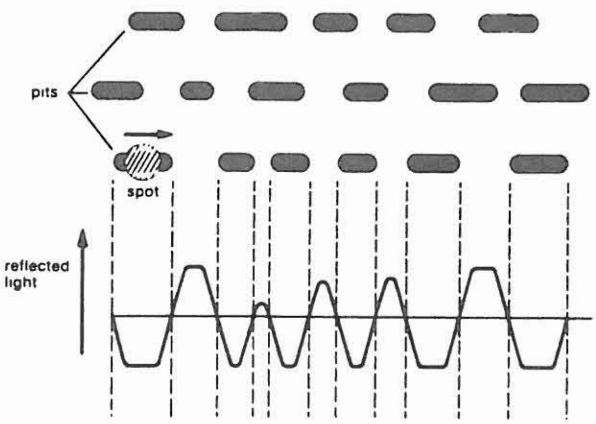
Signal processing

The information is present on the disc in the form of pits in a reflective surface. When the laser light hits a pit, the intensity of the reflected light returned to the photodiodes is much lower than of the light reflected back from the reflective surface between the pits.

In the Compact Disc player these differences in intensity are translated into an electric signal that may assume either of two levels: e.g. 0 volts if the intensity is high and + 5 volts if the intensity is low.
This signal contains all the information required to reconstruct the initial analogue audio signals.



How is this signal built up?
Before the analogue signal is stored on the disc, it has to be digitized.
At fixed intervals or time points the signal's amplitude is measured and the measured value is held for a short moment; this procedure is referred to as sample-and-hold method.
With the aid of an analogue-to-digital converter the amplitude of the sample is converted to a 16-bit binary value.
The 16 bit word is split into words of 8 bits called symbols.
This binary value is then further encoded by a rather complex sytem (see Appendix A) which allows the subsequent correction of errors which may occur during replay due to scratches, dirt, dust etc.
The encoding system also makes it possible to abtain 'sync' signals from the disc and to obtain a more favourable information density on the disc.
During readout of the information contained in the disc errors may be introduced which result from e.g. scratches, dust or dirt.
To have the possibility of correcting errors, if any, further symbols, called parity symbols,

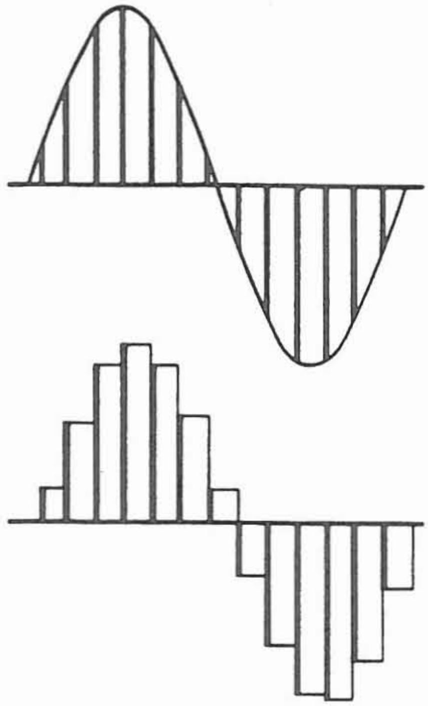


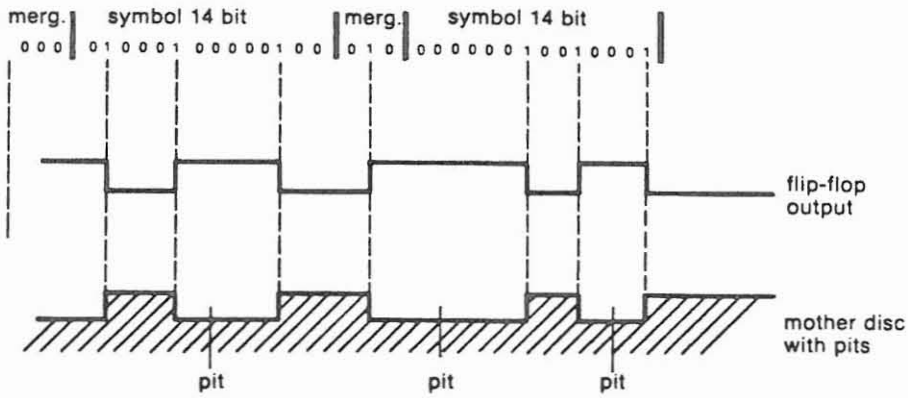
are added to the existing audio information symbols.

The string of 1s and 0s obtained by putting all the symbols after one another, is not suited to be recorded on the disc without more; these symbols are first fed to a modulator that converts them into 14-bit symbols.
These 14-bit symbols meet the requirement that a ONE is always followed by at least two ZEROS and maximum 10 ZEROS.

This requirement has been formulated to enable the player to obtain a 'sync' signal from the disc and a more favourable information density on the disc.
In between the symbols, the same requirement of minimum two and maximum ten zeros must be met for this, three additional bits are inserted in between the 14 bit code words, called merging bits.

The string of 0's and 1's obtained as a result of the above encoding system is then fed in serial form to a flip-flop.
The flip-flop changes states at every 1; its output signal is as shown in the Figure.
In turn, the flip-flop output drives a laser which burns the pits into the mother disc.
The results are also shown in the Figure.





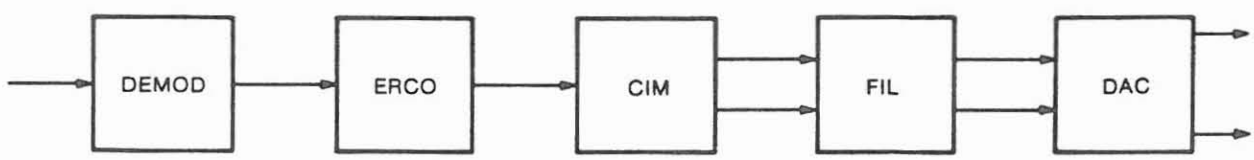
Reconstruction of the initial signal

The player has to reconstruct the initial signal from the one contained on the disc.
For this purpose, a sync signal, called clock signal, is first derived from the incoming signal.
This clock signal is used to recover the original 2x8-bit binary values. *Stereo*
This is achieved in the demodulator (DEMOD).
The 8 bit symbols leaving the demodulator are fed to the error-correction IC (ERCO).
This IC checks by means of parity bits whether errors have slipped into the symbols.
If this is the case, these errors are corrected.
The number of errors that can be corrected is, however, limited.

When the error-correction IC is incapable of correcting all of the errors, the information is passed on to of the next IC (CIM) in which the value of a sample can be interpolated from the values of the preceding sample and the next following sample.
If one of these samples is also missing, no correction is possible and the signal is MUTED.

From this interpolating and muting IC the sample values go to a digital filter (FIL) that allows the passage of the very restricted audio frequency band only and suppresses nearly all introduced frequencies which do not belong to the audio signal.

The signal leaving the digital filter is still a digital signal and has yet to be converted into an analogue signal. This is done in a digital-to-analogue converter (DAC).
At the DAC's exit we have the reconstructed initial audio signal which appears at the player's output after having passed a last filter.



Tracking

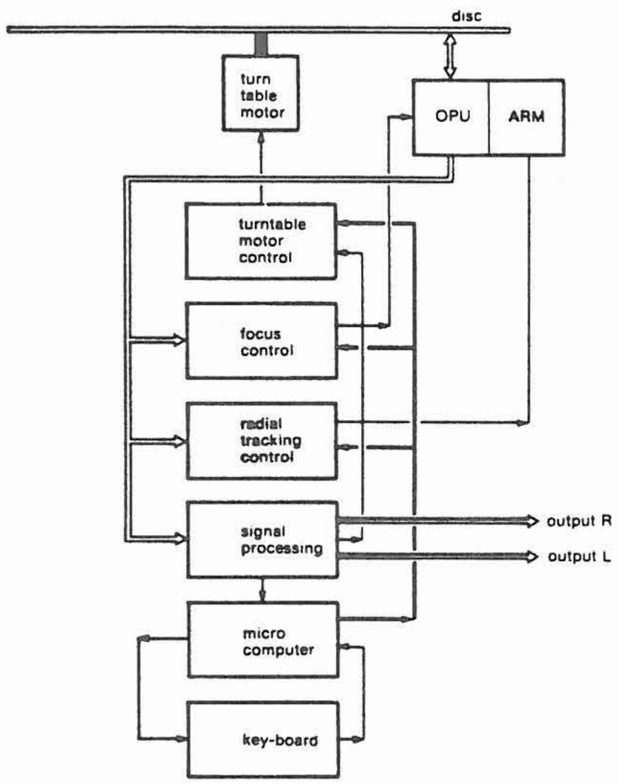
During disc readout the pick-up unit must follow the track with great precision to ensure a reliable signal.
Since any mechanical contact between pick-up unit and disc is absent, tracking must occur by electronic means. For this purpose reflected light is used.

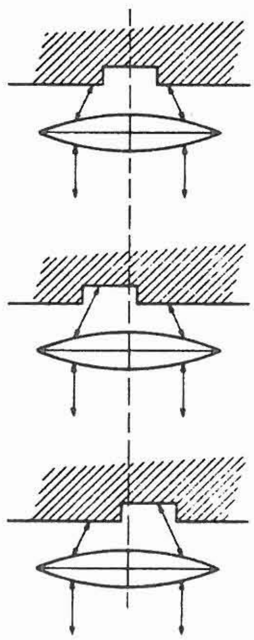
When the light beam is correctly centered on the track, the intensity of the reflected light will be equal at any point, that is, that amount of light reflected back from the right-hand side will equal that from the left-hand side. of the spot.

If the beam has shifted slightly and is partly falling e.g. at the right-hand side of the track, the amount of light reflected back from the right-hand side will exceed the amount returning from the left-hand side, because part of the light at the left-hand side is scattered by the track, whereas total amount of light is reflected back from the reflective disc surface at the right-hand side of the track.

A beam shift to the left will of course augment the amount of light reflected back from the left of the track so that it exceeds the amount of light returning from the right.

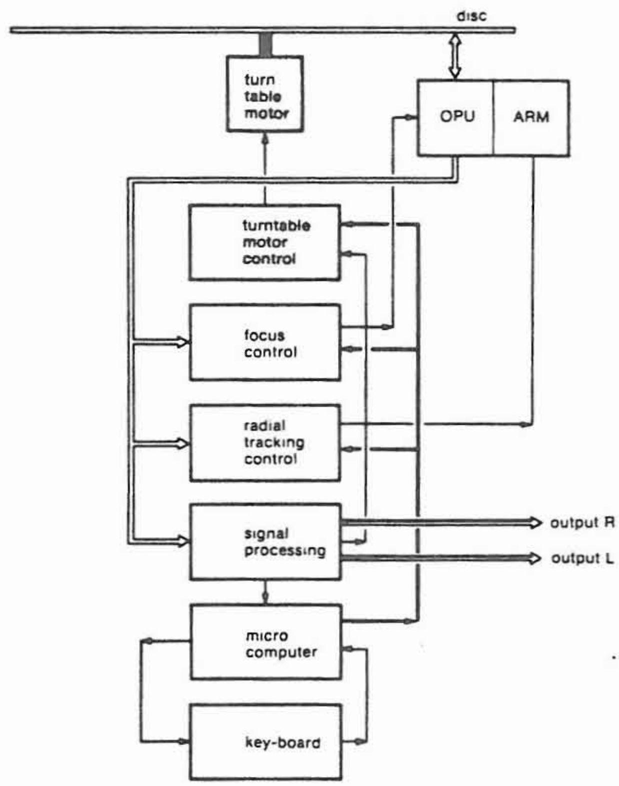
In the Compact Disc player the halves of the reflected light fall on two different diodes. If the amount of light on either diode is equal, the pick-up unit is accurately following the track; in the event of a deviation to the right or to the left, resulting in unequal amounts of light, a difference signal is obtained by subtracting the output signal of the diodes from one another. This difference signal is referred to as **radial error signal**.
This error signal is used to control the arm holding the pick-up unit until the error has been cancelled.





Turntable motor control

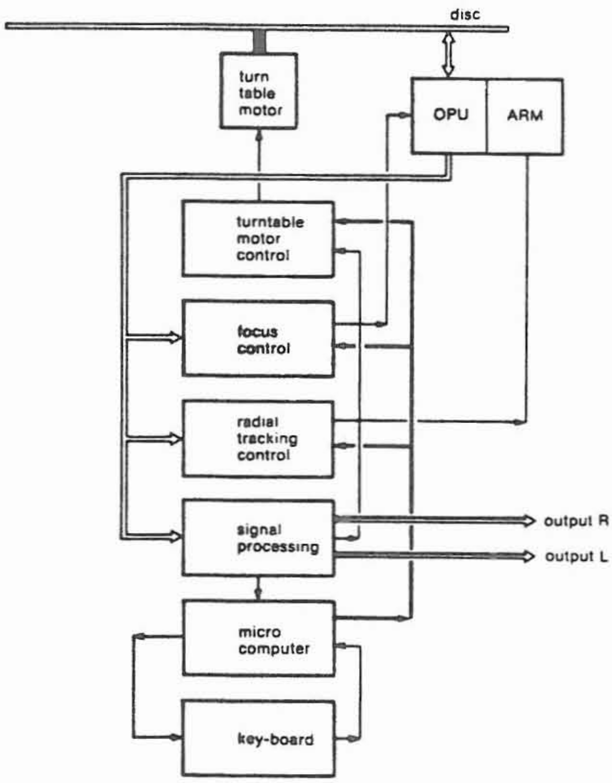
The flow of information or data must be maintained as constant as possible to the signal processing circuit.
The rate at which the bits enter is dependent upon the rotational speed of the turntable and upon the location of the pick-up unit relative to the disc i.e. near to or away from the centre.
The speed at which the track is moving with respect to the pick-up unit must be kept as constant as possible.
If the disc were rotated at a constant speed then when reading a track near the centre, the data rate would be lower than another reading of a track at the outer edge of the disc.
In the error-correction IC the rate at which information enters is compared with a constant frequency derived from a crystal oscillator. If too many bits per unit time are entering, the turntable motor is braked; the motor speed is increased when too few bits enter.
This maintains a constant data rate, and as a consequence, virtually eliminates wow and flutter.



Focus control

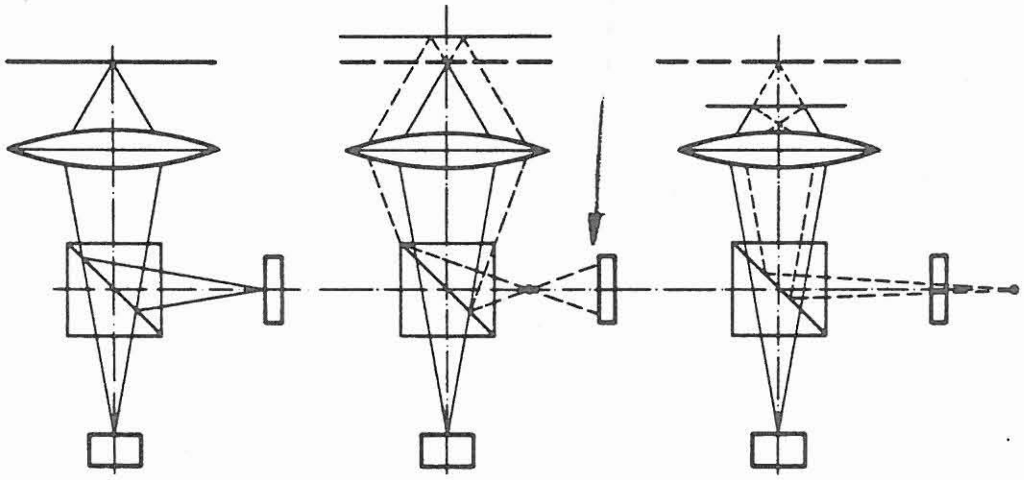
In order to obtain a reliable signal it is necessary that the light beam be properly focused on the track of the disc. Disc and turntable tolerances may cause variations in disc-to-objective spacing and, consequenlty, lead to incorrect focusing of the light beam on the disc, unless certain provisions are made.
To maintain a constant distance between disc and objective the objective is made movable in vertical direction.
By passing a greater or lesser amount of current in one direction through a coil, the objective can be moved e.g. down to a more or less extent.

By feeding the current in opposite direction through the coil an upward movement is generated.
Any focus error is detected by a difference signal derived from reflected light on the photodiodes and this signal is used to control the focus motor via a servo system controlled by a microcomputer.



The information whether the distance is too great, too small or perfect can be obtained frc the reflected light.
If the focal point coincides with the information plane on the disc, the focal points of the reflected light will also coincide with the diodes. Too great or too small a distance between disc and objective will bring the system out of focus and the focal points of the reflected light are in front of or behind the diodes.

The special arrangement of these diodes allows to determine whether or not the system is perfectly focused and to correct, if necessary, the objective-to-disc distance.



The microprocessor

The various different functions in the player need be correctly integrated (to one another), and also the commands entered via the keyboard must be processed.
A microcomputer is used to achieve this.
It checks the various functions for correct operation and ensures that - e.g. during start-up of the player - the functions are correctly sequenced.

The conditions which the microcomputer monitors are:

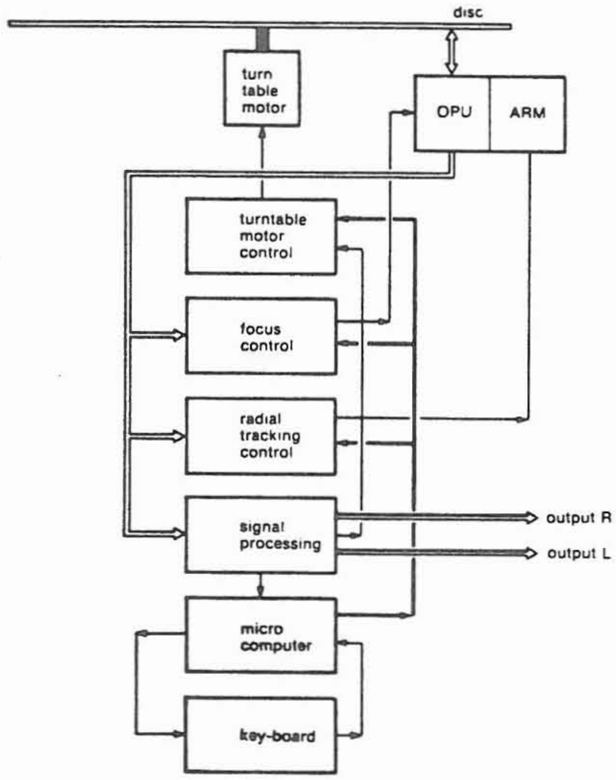
Is the system perfectly focused?

Does there exist a data flow?

Is the track correctly followed?

Has a key been depressed on the keyboard?

All these things are checked and, in the event of a deviation signals are generated to cause the equipment to respond appropriately.
The cycle of the microprocessor contains a fixed programme for keyboard scanning; this programme is carried out at fixed time intervals.
If a key is depressed, the microprocessor sees to it that the command is carried out by instructing the related circuits.



Summary

In the Compact Disc system the information is contained in a reflective disc which is read out by means of an optical system.
The information present in digital form, is pressed into the reflective surface in the form of pits or depressions.
Inside the player the information is de- modulated and existing errors are corrected before a digital-to-analogue converter reconverts the digital signal into the initial analogue audio signal.

The optical system is composed of a laser, a set of lenses and light-sentitive diodes.
The lens system ensures correct focusing of the laser light on the disc surface, whereas the photodiodes convert the reflected light into electrical signals.

Light that hits these pits is scattered in such a way that the intensity of the reflected light falling on the photodiodes is much lower than that of the light reflected back from the reflective surface between the pits.

During disc read-out the pick-up unit will have to follow the track with great accuracy.
Using more than one photodiode makes it possible to measure the deviation and hence correct the position of the arm holding the pick-up unit.

The speed at which the turntable is rotating also needs regular adjustment.
To achieve this, the rate at which the information enters is measured.

The accuracy of focusing can also be determined by means of the reflected light.
If the system is out of focus, correction is achieved by controlling the objective.

A microcomputer controls the operation of the player and reads the keyboard, thus integrating all the functions.

The Compact Disc Player

Questions

The disc

1. *What is the disc composed of?*
2. *At what speed is the information track scanned?*
3. *At what speed does the disc rotate?*

The optical pick-up unit

1. *What kind of laser is applied in the CD player?*
2. *How many light beams emerge from the laser?*
3. *On what is the amount of emergent light dependent?*
4. *How is the intensity controlled?*
5. *What is the function of the*
   . prism?
   . photo diodes?
6. *In what way is a split-up of the reflected light beam into two equal beams obtained?*
7. *How is the objective moved up and down?*

Description of the decoder

1. *What kind of output signals are supplied by the demodulator?*
2. *How is the bit clock re-generated from the incoming data?*
3. *How can DEMOD determine where the flow of data begins?*
4. *What does modulation imply?*
5. *Which tasks are performed by the error corrector IC?*
6. *How is the motor speed controlled?*
7. *How is de-interleaving obtained?*

8. What happens in the error correction decoders in case of
   - no errors?
   - one error?
   - two errors?
   - more than two errors?
9. What function has the interpolating and muting IC?
10. What processing is first performed on the data in the CIM?
11. Which actions are undertaken by the CIM if
   - a single error is flagged?
   - two or more errors are flagged?
   - after a drop-out a new signal is present?
12. What filtering system is used in the Philips Compact Disc player?
13. Which method is used in this filter?
14. How many bits can the DAC convert to an analogue signal?

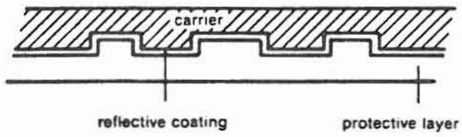
Servo systems

1. Which mechanical movements in the CD player are controlled by a servo system?
2. When will a servo system become unstable?
3. How is a servo system in the Compact Disc prevented from becoming unstable?
4. How is the stability of a system affected by the gain?
5. How is the radial error signal obtained?
6. Which other factors can make the radial tracking system unstable?
7. Why have the factors d and k been introduced?
8. How are the factors d and k obtained?
9. How are the signals of the photodiodes affected by a change in the distance between disc and objective?

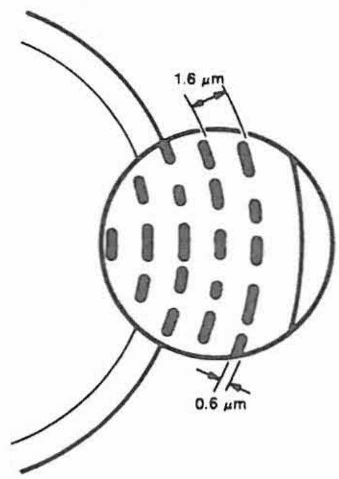
The Compact Disc player

The disc

This chapter will enter into more detail on the Compact Disc player, starting with the disc and the optical element and followed by the signal path and the servo systems.



The CD disc is composed of a plastic base material or carrier into which the information has been pressed in the form of pits. Next, a reflective layer is placed over the pits and a transparent protective coating is added over the reflective layer.
The total playing time of a CD disc is 60 minutes.
The diameter of the disc is 120 mm.
The average pitch of the information track is 1.6 μm; the width of the pits is 0.6 μm.
The disc is single sided.
The information track is scanned at a constant speed of 1.3 m/s. To achieve this, the rotational speed is gradually stepped down from 500 rpm at the beginning of the disc near the centre of the disc to 200 rpm near the outer edge of the disc.
(⇒ constant data flow)

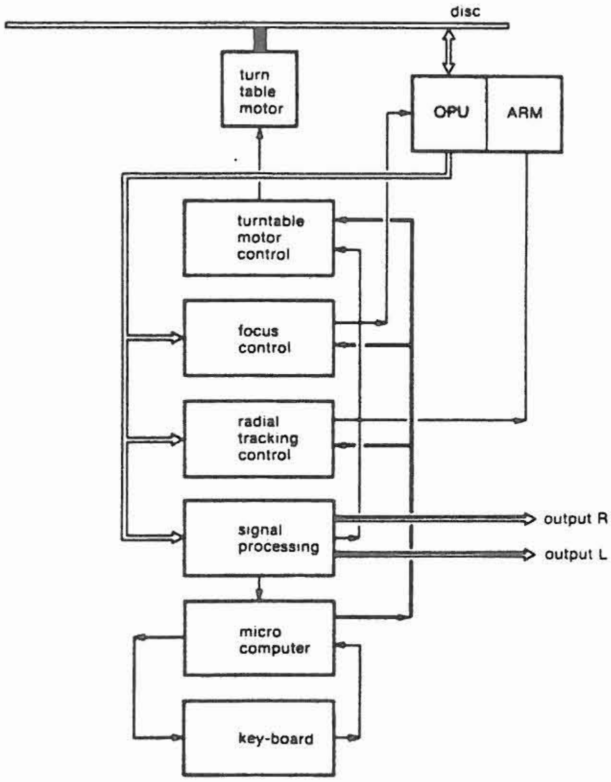


The optical pick-up unit

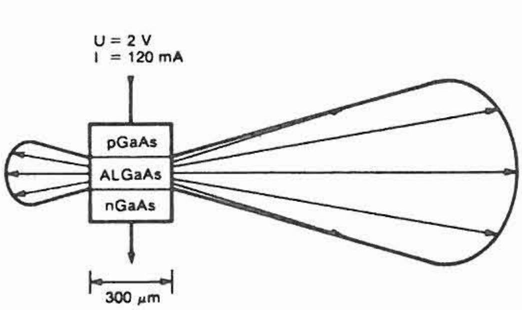
The optical pick-up unit (OPU) is composed of:

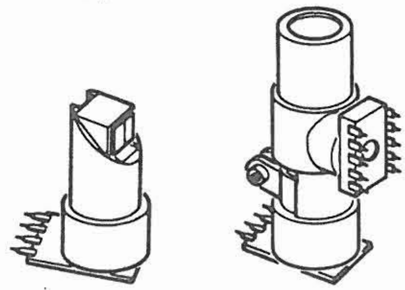
- the laser unit
- the collimator
- the photodiodes
- the focusing unit

The laser and the prism are accommodated in the laser unit; the collimator and the photodiodes have also been combined into a single unit.
The focusing unit comprises the objective and the mechanism that allows to move the objective up and down.



The laser
The laser applied in the CD player is a semiconductor laser.
This laser generates two light beams, that is, a main beam emerging from one end and a secondary beam emerging from the opposite end.
The main beam goes via the prism and the set of lenses to the disc; the secondary beam falls on a photodiode that has been incorporated in the laser casing.
There exists a fixed relationship between the intensities of the main and secondary light beams. An important factor for good operation of the player is that the amount of light falling on the disc is kept as constant as possible with time. The intensity is, however, dependent on the current fed through the laser and on the temperature of the diode. Consequently, the intensity must be kept constant by controlling the laser current. The light of the secondary beam is taken as measure for the intensity of the main beam.

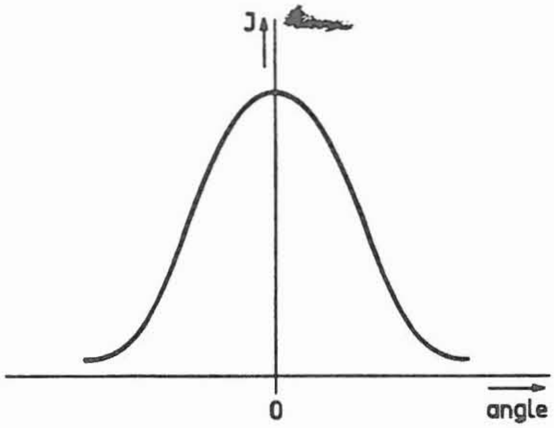




The amount of light of the secondary beam is measured with a photodiode installed below the laser. The signal of this diode goes to the laser supply and thus determines the current through the laser and, consequently, the intensity.
If the instensity is too low, the diode signal is too small and the current will be increased; if the intensity is too high, the diode signal will be too strong and the current through the laser is reduced.

A bit of theory on laser light

The light intensity pattern yielded by a laser diode as a function of the angle is shown in the Figure.



This radiation pattern is called the far-field intensity pattern.
The cross-sectional form of the spot
[light intensity = f (cross section)]

The Term FWHM stands for Full Width Half Maximum, that is, Full Width at half the maximum value (= over-all half-value).

The total half-value for the Philips laser diodes is:
parallel to the laser approximately 30°
perpendicular to the laser approximately 50°.

The spot on the disc applies:

$$\varnothing = \frac{0{,}6\ \gamma}{NA}$$

where:
$\varnothing$ = the size of the spot (= FWHM)
γ = the wavelenght of the light
NA = the numerical aperture
(NA is the sine value of the half aperture angle).

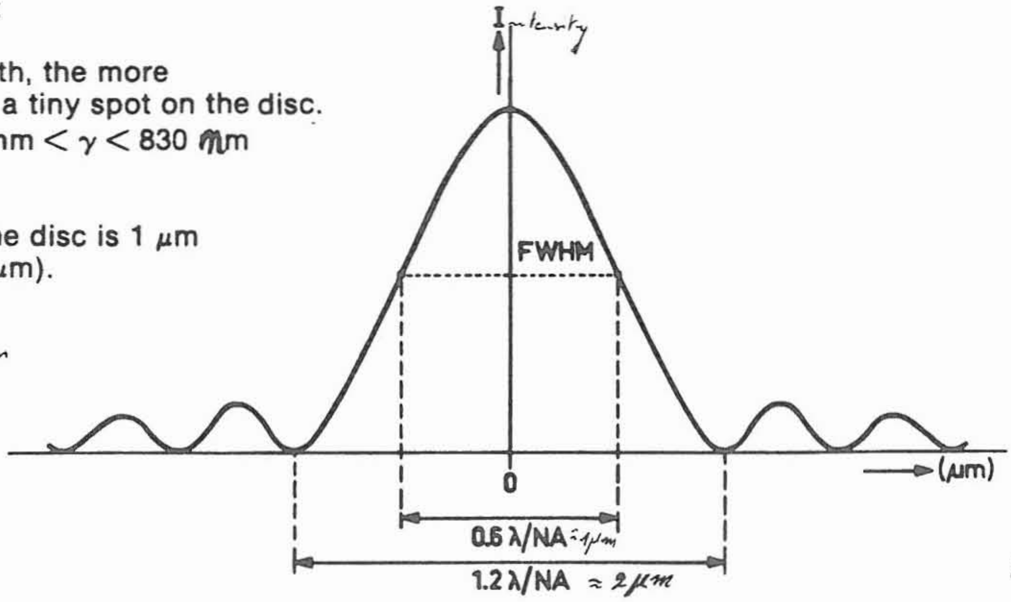
From the formula follows:

The greater the wavelength, the more difficult the realisation of a tiny spot on the disc.
To CD applies: γ : 780 nm $< \gamma <$ 830 nm
NA: 0,45

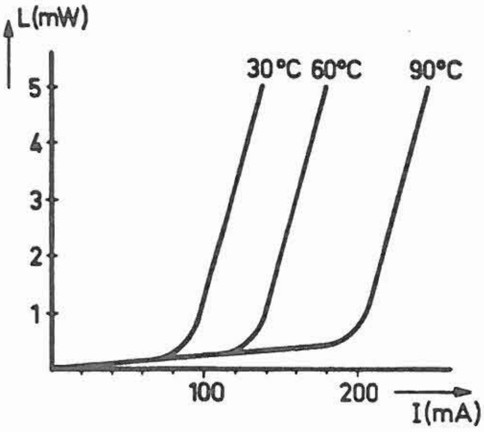
The size of the spot on the disc is 1 μm (depth of focus: about 2 μm).

γ = 800 nm

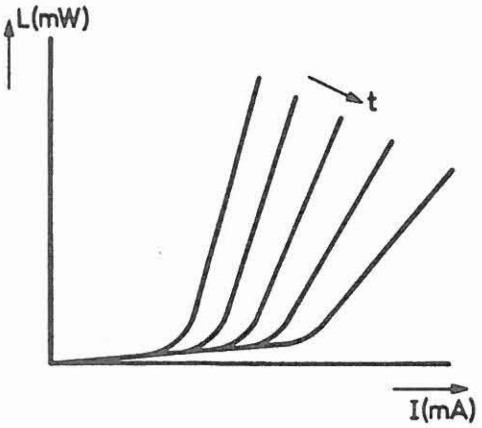
$\varnothing = \frac{0{,}6\ \gamma}{NA}$ = 1066 nm
= 1,066 μm



The lifetime of a laser diode is directly coupled to the temperature.
The figure shows the temperature dependence of the light current characteristic.



The graph illustrates that the treshold current increases at increasing temperature.
The deterioration is shown in the next Figure.
With time, the characteristic becomes less steep and the threshold current intensity becomes ever stronger.
When the light intensity or the heat development in the laser becomes too great, this phenomenon takes the form of an instantaneous damage.
When appropriate measures are taken, the effect will hardly be perceptible after thousands of burning hours.
(This stresses the importance of a transient suppression system.)



Laser light to the disc

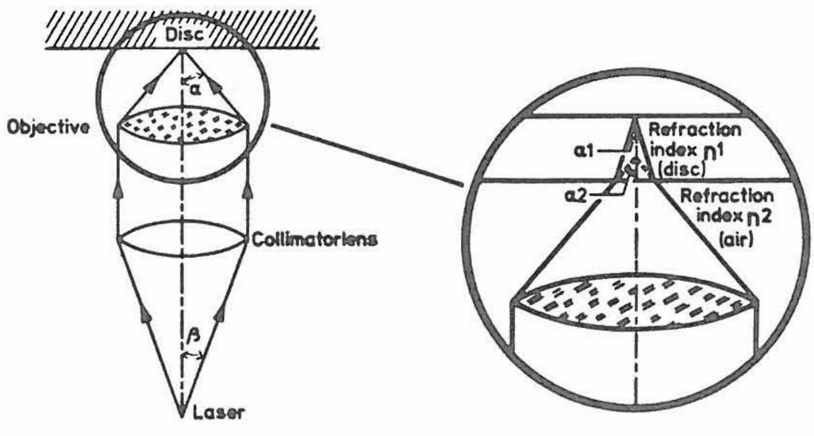
The diverging laser beam should reach the disc in the form of a spot.
This is achieved by means of a set of lenses.

The spot diameter on the disc is 1 μm.
The laser output is 1,5 mW $\div$ 3 mW.

The luminous efficacy on the disc is 0,1 mW.
Numerical Aperture (NA):

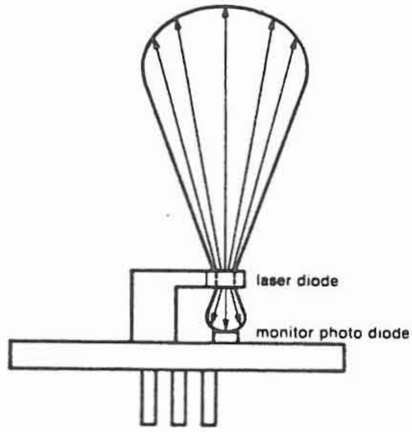
at NA $\alpha = 0{,}45$ (α approx. 27°)
NA $\beta = 0{,}11$ (β approx. 6,5°)

NA $= n_1 \sin \alpha_1 = n_2 \sin \alpha_2$



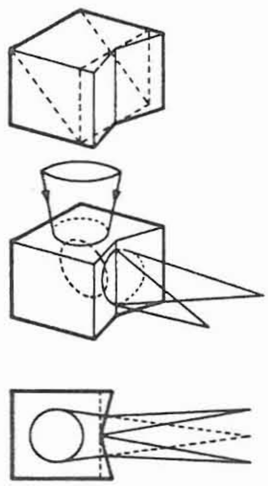
Prism and photodiodes

The light that leaves the laser goes to a semi-reflecting prism; this prism lets pass the light coming from the laser and deflects the light beam returning from the disc to the photodiodes.
The photodiodes convert the reflected light into electric signals. A total of four diodes have been mounted slightly apart in groups of two.
The light belonging to one half of the reflected light beam falls on one diode pair, that belonging to the other falls on the other diode pair. To obtain a split-up of the light beam into two equal beams, one side of the prism has been ground to wedge shape.



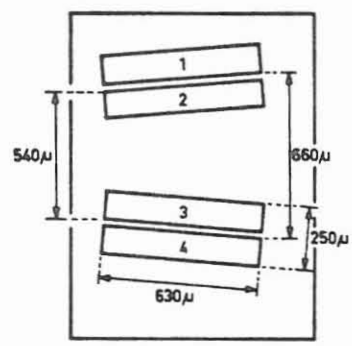
The collimator

After the laser beam has passed the prism, it goes to the collimator where the diverging beam is transformed into a parallel beam.
The advantage of a parallel beam is that the distance between collimator and the next lens does not influence the functioning of the set of lenses.
This feature is required because the objective that follows the collimator has to move in a vertical direction to maintain perfect focus with a disc moving in a vertical direction due to tolerances.

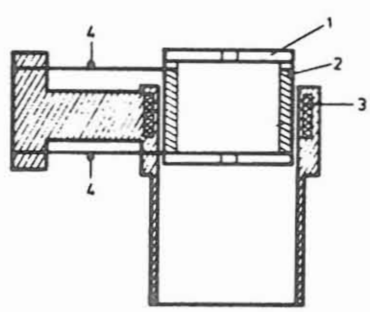


The objective

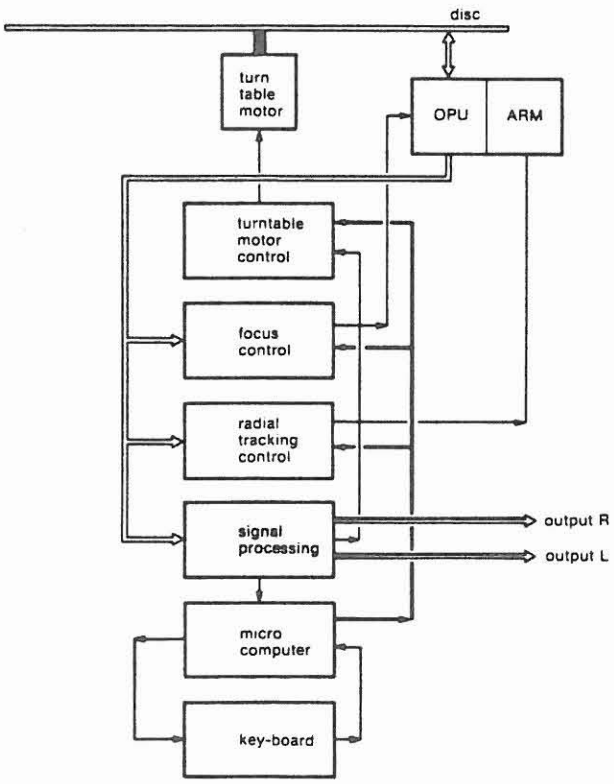
The objective is installed on a ring of magnetic material; the latter is attached to a housing by means of two leaf springs.
This housing contains a coil which produces a magnetic field when an electric current is fed through it.
This magnetic field being proportional to the current traversing the coil, exerts a force on the magnetic material onto which the objective has been mounted, thus making the objective move up or down, depending on the polarity and magnitude of the coil current.



1 objective
2 magnet
3 coil
4 leaf spring



Description of the decoder



Introduction

The task of the decoding system within a CD player is to re-generate both analogue audio channels from the high-frequency data being retrieved from the disc by the optical pick-up unit. The block diagram of a decoder is shown in the Figure where emphasis is put on the interface lines to and from the servo systems.

Special LSI circuits have been developed for the signal processing tasks of demodulation (DEMOD) error detection/correction (ERCO) and interpolation/muting (CIM).
The digital-to-analogue conversion is based upon a dedicated digital filter (FIL) and a 14-bit D/A converter (DAC) which, when combined, give the performance of a 16-bit system with the additional benefits of a linear phase characteristic and reduced distortion.

When the CD is used for reproduction of sound only, as with conventional players, the high frequency detector and subcoding processor display unit are not required. However, their inclusion allows certain additional sophisticated functions to be realised.

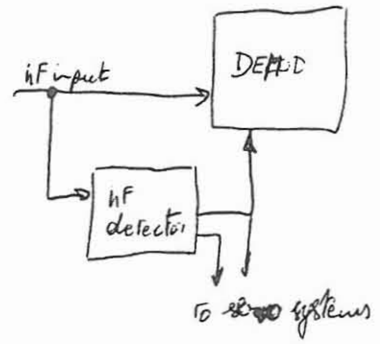
DEMOD, the demodulator

The demodulator IC forms the front-end of the CD decoding system and supplies demodulated data and timing signals to the error corrector IC and the subcoding microcomputer. A simplified block diagram of the DEMOD is shown in the Figure.

The high frequency (HF) signal retrieved from the disc by means of the optical pick-up unit is amplified and filtered externally to the decoder and then supplied to the input of de demodulator.

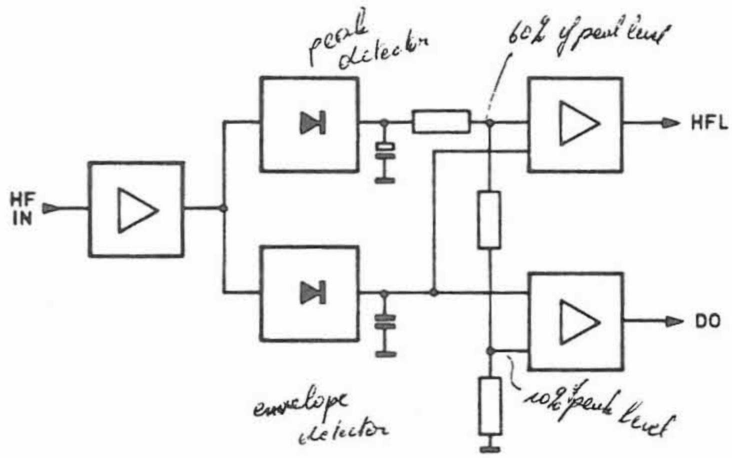
The HF detector

The use of a detector for the high-frequency input signal is optional, i.e. the HFD input at DEMOD may simply be conneted to the supply voltage to enable normal operation of the demodulator IC. Using a HF detector can, however, improve the performance of the whole CD player. A proposal for such a HF detector circuit is shown in the Figure.



Zie ook blz 29 in schema decoder (links boven)

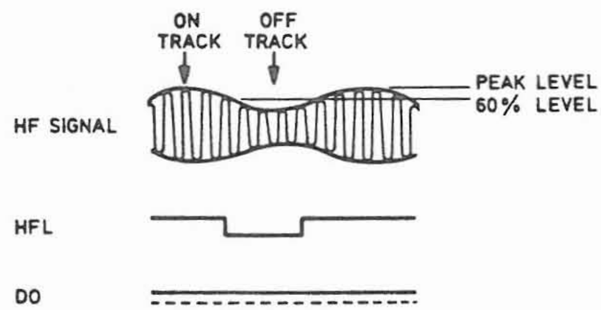
The HF input signal is first amplified and then supplied with opposite phases to a clamp circuit and two full-wave rectifiers which operate on filter circuits with largely differing time constants.



There are two output signals generated in this HF detector circuit, HFL (high-frequency low) and DO (drop-out). Both of them are routed to the servo systems for further processing.

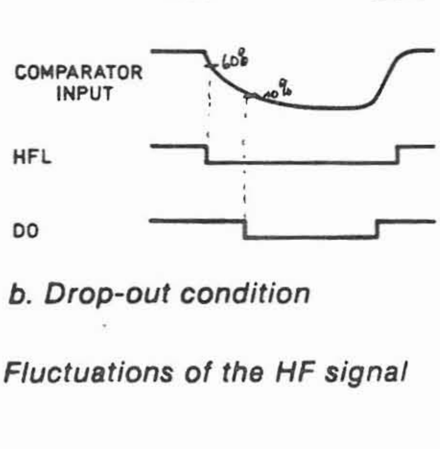
The principle of operation of this HF detector relies on the different time-constants of the two rectifier/filter circuits. One of them operates as a peak detector with a rather long time-constant of about 100 milliseconds, while the other one tracks the envelope of the HF signal with a time-constant of about 20 microseconds.

A comparator circuit provides switching levels of approx. 60% and 10% of the peak value of the HF signal. This is used for the analysis of fluctuations of the amplitude of the HF input signal. For the following two paragraphs, please refer to the next Figure.



a. Loss-of-track condition

Assume a small radial displacement of the moving arm which carries the optical pick-up unit in a CD player, e.g. due to vibration or a shock. In that case, the amount of light which is reflected from the disk begins to decrease and, therefore, the amplitude of the HF signal decreases, too. When the amplitude falls below about 60% of the initial peak value, the upper comparator in the Figure switches and HFL goes low. This is shown in the Figure. As the lightspot used for scanning the disk has finite dimensions, the amplitude of the HF signal does not reach zero between tracks. Therefore, the higher of the two switching levels is used for this kind of signal fluctuation.



b. Drop-out condition

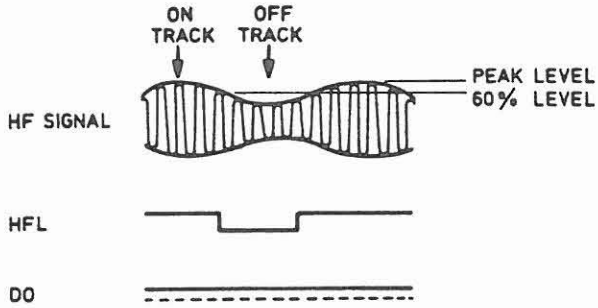
Fluctuations of the HF signal

In case of a drop-out, e.g. due to dust or scratches on the disk, the HF signal can remain at full (all light reflected) or zero level (no light reflected).
In that case, no AC input signal is generated for the rectifier circuits and the output

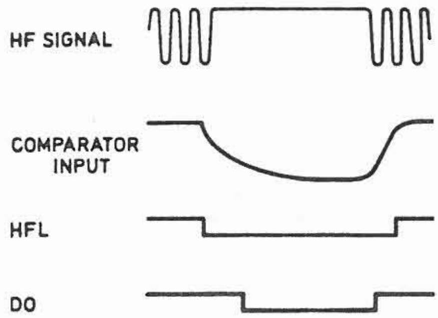
→ envelope detector.

/oltage of the fast filter circuit decays ɜcc⌐rding to the time-constant of that filter as ;hc , in the Figure. When a level of approx. ;0% of the initial peak amplitude is reached, -IFL goes low. When a level of about 10% is 'eached, DO goes low. By interpreting the ;ignals HFL and DO, the servo processor can listinguish between a loss-of-track situation correction of the position of the lightpen is 'equired) and drop-outs (no correction 'equired).

This HF detector circuit improves the ɔerformance of a CD player, because it ;enses the level of the HF input signal and lisables the phase and frequency detectors n DEMOD in case of insufficient amplitude. Thereby, the PLL cannot lock onto noise in he absence of a HF signal and clock jitter is ɪvoided. Further, fluctuation of the amplitude ɔf the HF input signal are interpreted so that o f track and drop-out conditions can be ɔgnised separately and the servo ɔr ɜssor may act accordingly. Moreover, the -IFL output may be used by the servo ɔrocessor for the implementation of search unctions or other sophisticated operations.

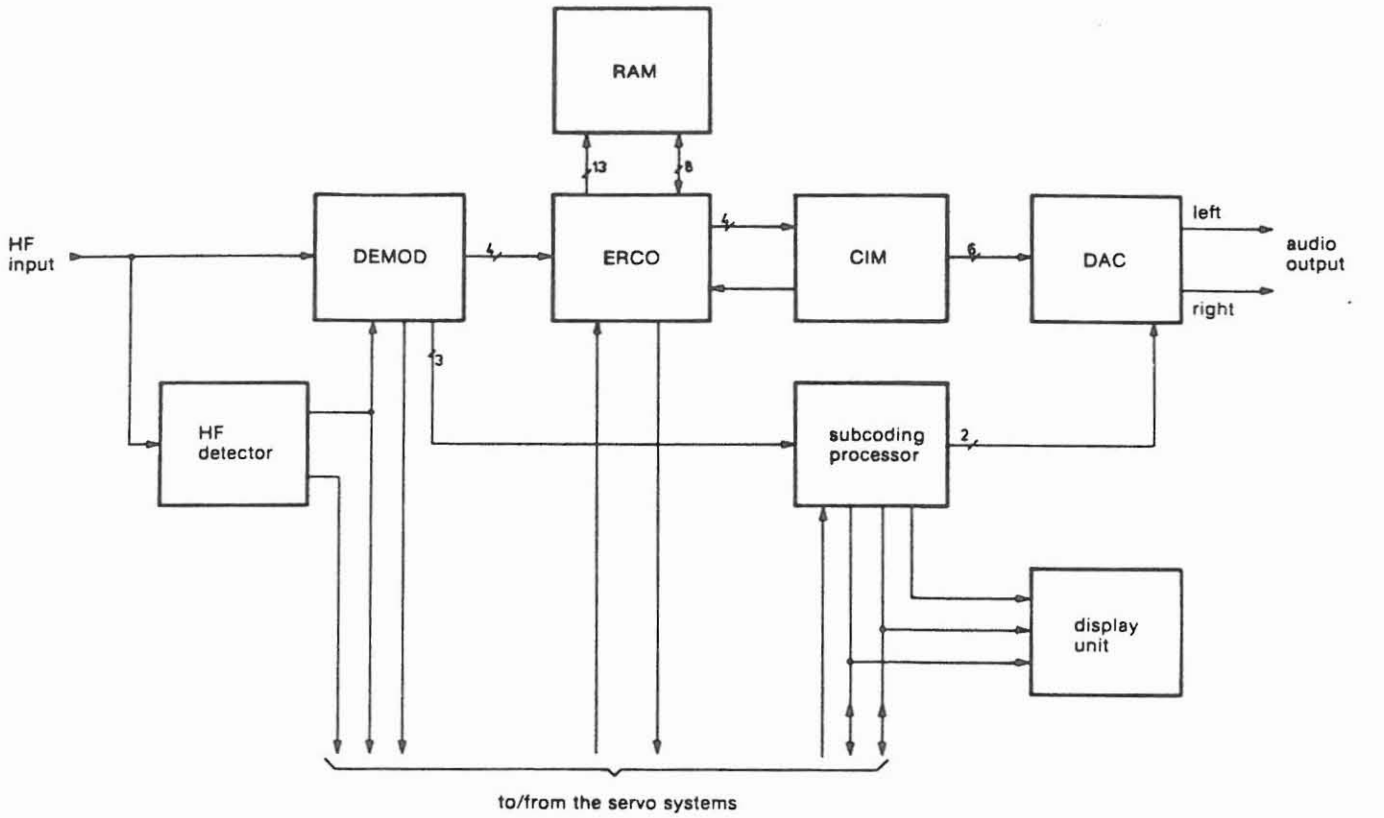


a. Loss-of-track condition

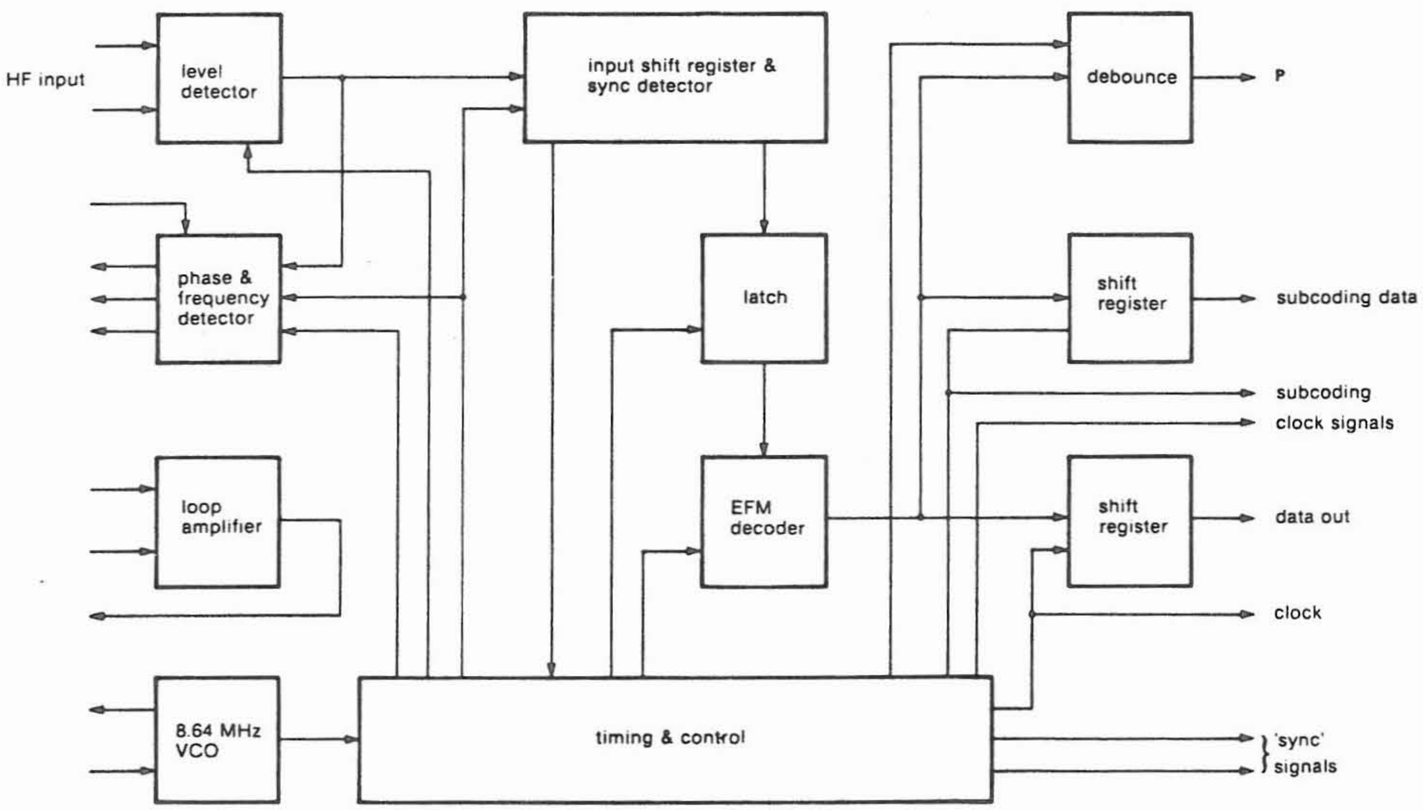


b. Drop-out condition

Fluctuations of the HF signal



DECODER



DEMOD

The analogue HF signal is first digitized and then passed to a phase-locked loop (PLL) which re-generates the bit clock from the incoming data.

There are six complete stereo audio samples in one frame of disc data. As each frame contains 588 channel bits, it can be computed that one stereo sample results in 98 channel bits.
If we multiply this figure by the sampling frequency of 44.1 kHz, we get the data rate of the HF signal which is nominally 4.3218 Mbits/s.

The voltage-controlled oscillator (VCO) of the PLL generates a. 4.3218 MHz master clock signal from which all internal timing is derived.

The incoming data is received in an input shift register. Provision is also made for the detection of the sync pattern which indicates the beginning of a frame of data.
This information is passed to the timing and control logic in order to synchronize the demodulation process to the received data.

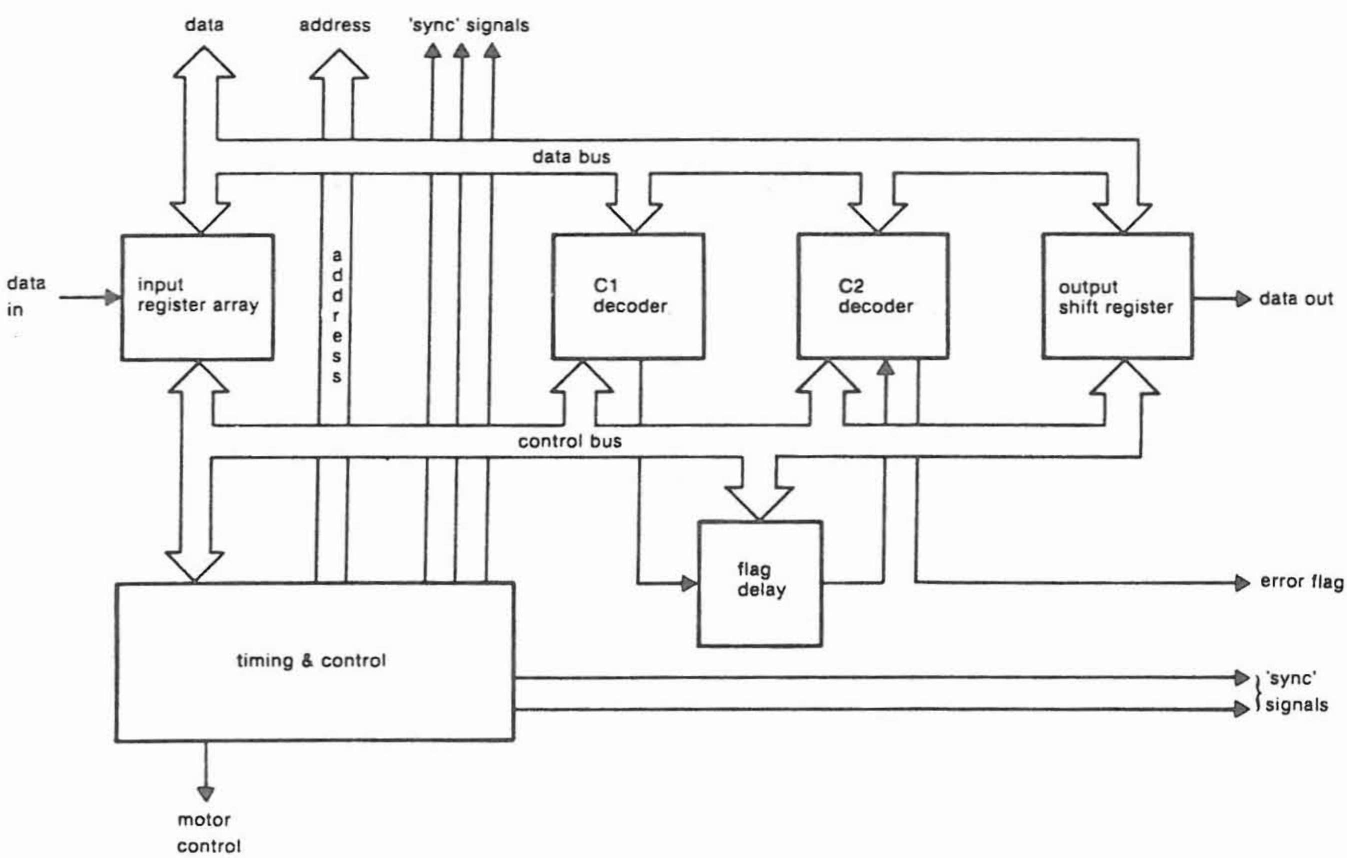
Provided the timing generator is locked to the HF signal, each received 14-bit word is stored in a latch and is then demodulated into an 8-bit symbol by the EFM decoder.

The demodulated audio data are shifted out to the error corrector with the timing and control logic providing the necessary clock signal and the symbol and frame sync signals.

E[illegible]O, the error corrector IC
The error corrector IC performs the tasks of error detection and error correction and supplies data to the interpolation and muting IC together with a flag signal indicating whether a concealment action has to be started or not.
A simplified block diagram of the error corrector is shown in the next Figure.

Data from the demodulator IC is organized in frames of 32 symbols of 8 bits each. Of these symbols 24 are audio symbols (one symbol is half an audio sample of 16-bits) and 8 are parity symbols added for error detection and correction purposes.

Data is entered serially into a register array. This array consists of a shift register which accumulates the symbols for parallel processing and a first-in, first-out register (FIFO).
This FIFO acts as a jitter reduction circuit, i.e. a deviation from the nominal data rate of up to $\pm$ 2.25 frames of data can be compensated for. It is due to this de-jitter operation that effects like wow and flutter are eliminated.
The speed of the output data of ERCO depends only on a clock signal which is derived from a crystal oscillator.
From the difference between the nominal and the actual data rate an error signal is derived for the control of the motor speed. This is a pulsewidth-modulated signal with a range of 142 linear steps.



ERCO

'he interleaving of data symbols (being a ıart of the CIRC error correction encoding ıroc ʿure, see Appendix A) is compensated or by suitable de-interleaving operations in ıe error corrector.
'he de-interleaving is performed on data ymbols prior to entering the C1 and C2 lecoders by temporarily storing these ymbols in an external RAM.
'he data symbols from the demodulator are le-interleaved by means of the RAM and the 2 symbols forming the input word to the C1 lecoder are checked for errors.
3y multiplying the input word by the parity :heck matrix, four syndromes are obtained. If here are no errors then all four syndromes ıre equal to zero and the 28 data symbols at he output of the C1 decoder (4 of the parity ;ymbols are discarded here) are written back nto the RAM unchanged.
n case of one erroneous symbol this symbol ... ected and the 28 corrected output ...bols are written into the RAM. In case of w... or more erroneous symbols the 28 output ymbols are written into the RAM unchanged ınd a flag is set to mark these 28 symbols as ɔeing unreliable.

f an erroneous C1 word results in a C1 flag ıen all of the symbols in the C1 output word re unreliable.
ıs all of these symbols have a different delay ıefore reaching the C2 decoder, the C1 rasure flag also must have different delays or the various symbols in such a way that an rasure symbol and its accompanying erasure lag arrive at the input of the C2 decoder at he same time.

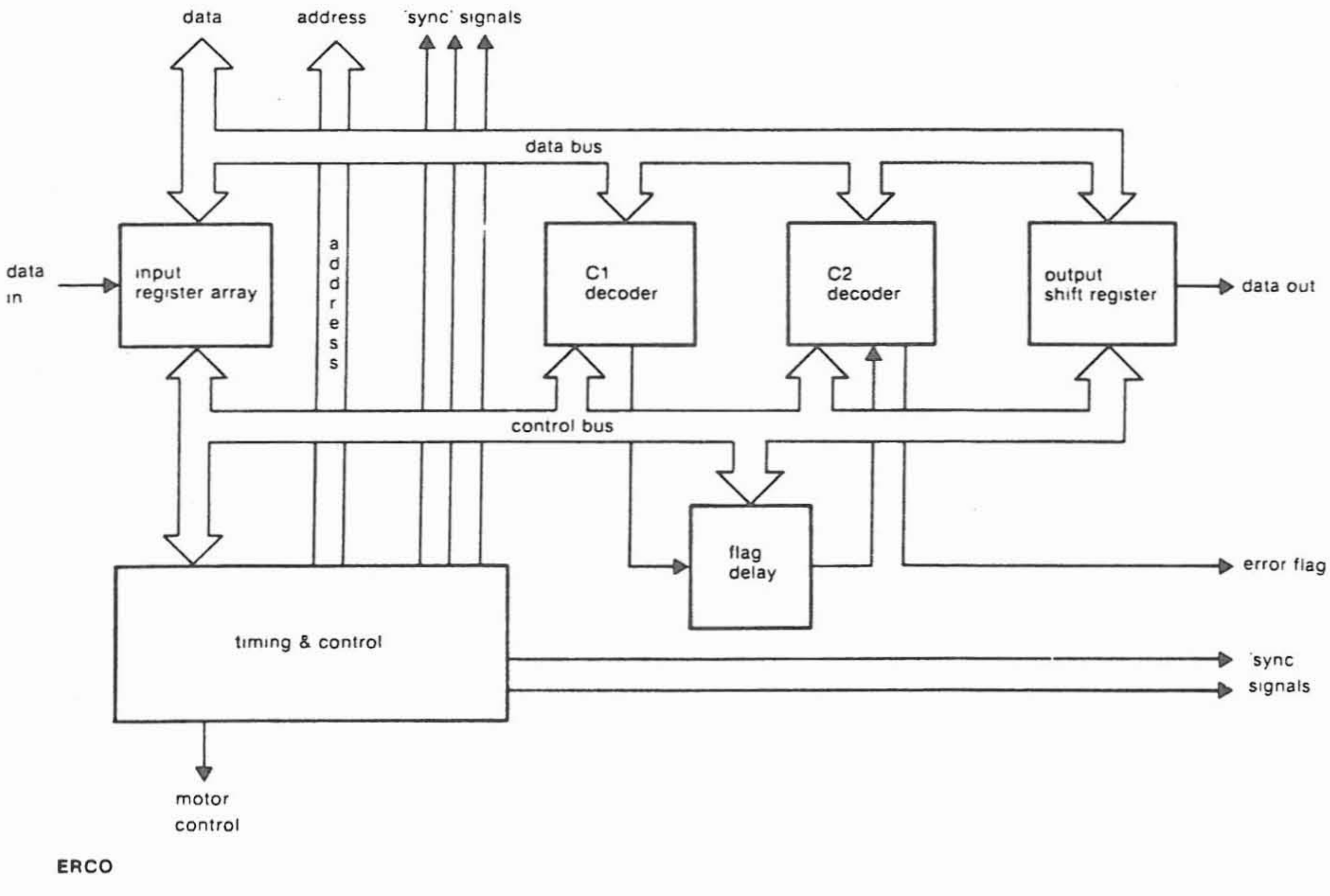
'he output symbols of the C1 decoder are ... erleaved by means of the RAM and the ... symbols forming the input word of the C2 le... der are also checked for errors by ːxamining the four syndromes that result rom the multiplication with the C2 parity :heck matrix.
f there are no errors then all four syndromes ıre equal to zero and the 24 data symbols at he output of the C2 decoder are written back nto the RAM unchanged; the remaining 4 ɔarity symbols are discarded here.

n case of one erroneous symbol this symbol s corrected in the same way as described for he C1 decoder.
The corrected 24 output symbols are written nto the RAM. In case of two erroneous ;ymbols the sum of the error values is given ɔy one of the syndromes.

Information about the position of the erasure symbols is available due to the erasure flags. With this additional information both errors are corrected and the 24 corrected ouput symbols are written into the RAM.
In case of more than two erroneous symbols in a C2 word all 24 symbols are written back into the RAM unchanged and a C2 flag is set to mark these 24 symbols as being unreliable.

The unreliable data signal provides two different kinds of information. If the flag is output simultaneously with a data symbol it marks that symbol. If it is output in the gap between two data samples it marks that sample that will reach the output 60 sample times later. This advanced flag is needed for the interpolation and muting circuit.

The output data are shifted out serially together with the data clock and the frame sync signal.

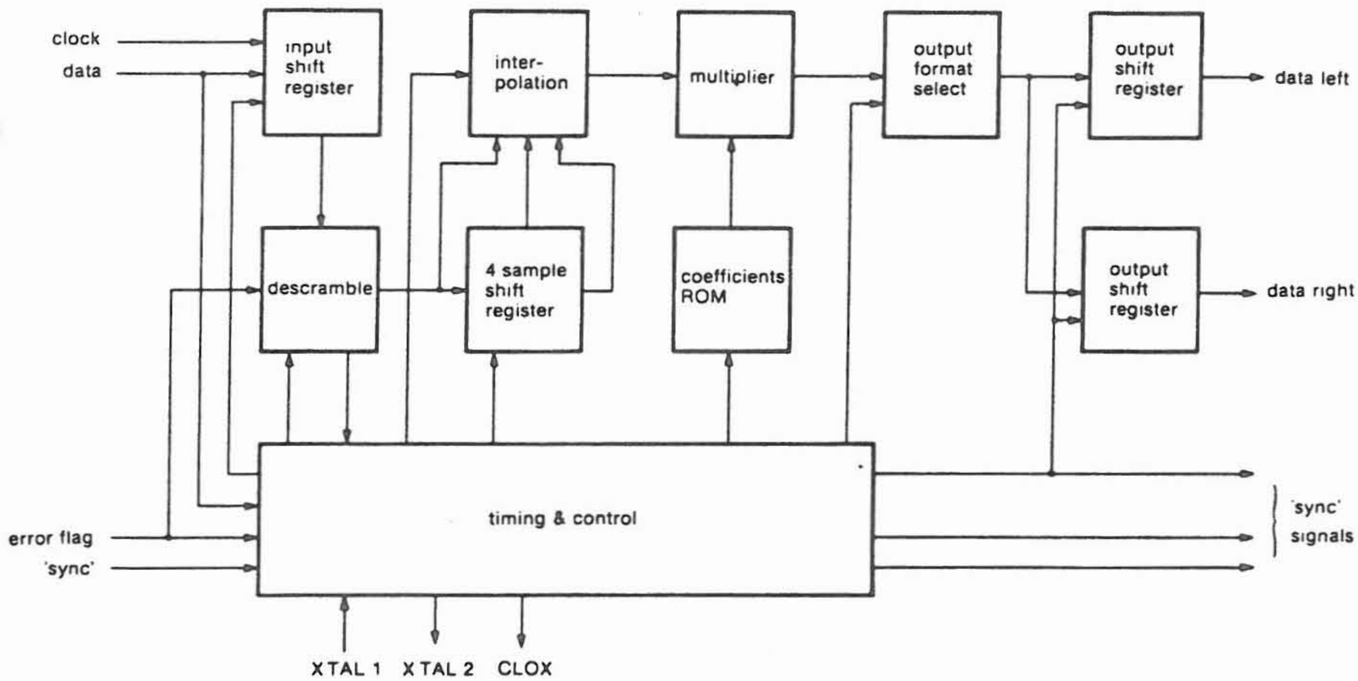


ERCO

CIM, the interpolation and muting IC
The interpolation and muting IC is used to reduce audible annoyance that could be caused by uncorrected erroneous data samples from the error corrector IC. A simplified block diagram is shown in the Figure.

Data from ERCO is entered serially with the shift clock.
After that, data are descrambled (see Appendix A) and separated into left and right channel samples.
performed on the unreliable data flag. When no error flags occur, the data value is not affected through CIM.

If for either left or right channel a single 'error' is flagged between two 'good' samples then linear interpolation is used to replace the erroneous value. If two or more adjacent samples are flagged then the samples in error will be muted.
Beginning 30 samples before the consecutive errors, the value of the samples will be attenuated smoothly to zero level, using a (0 to π) cosine-shaped curve.
The coefficients of that curve are stored in a ROM and the attenuation is performed by a multiplication of the sample values with these coefficients.



CIM

After the error burst the next 30 samples are processed with a (π to 2 π) cosine-shaped curve after which the full level is reached again. This muting or attenuation occurs to data of the left and right channel simultaneously regardless of the source of the mute.

A crystal oscillator is used to generate the internal timing signals as well as the clock signal CLOX for the error corrector and the digital filter.
By using the frame sync signal for internal timing reset, CIM automatically synchronizes to the ERCO output.

Digital-to-Analogue Conversion
The last section in the sequence of processings to which the signal in the Compact Disc system is subjected, is the re-transformation of the digital code into an analogue signal that has the same shape as the original acoustic vibration that was captured by the microphone.

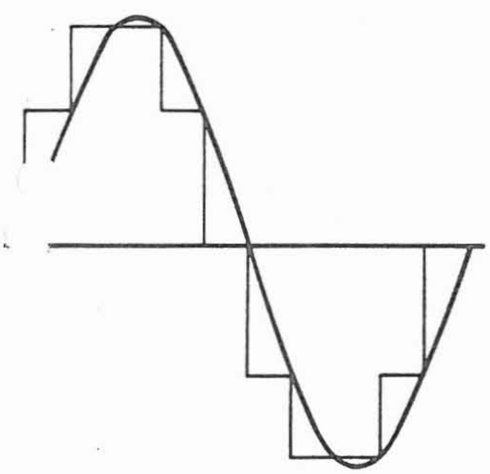
The digital signal takes - after decoding and error correction - the form of a series of 16-bit words.
Each word represents in binary form the numerical quantity of the sound magnitude at that moment and is, therefore, a sample of the audio signal. Each second 44100 of these words appear.

The Digital-to-Analogue Converter (DAC) in the Compact Disc player generates with each word an electric current of associated strength and keeps it constant untill the next word arrives. The electric current will thus describe a 'stepped' curve that is an approximation to the analogue signal shape. However, a step function contains harmonics which far exceed the band width of the analogue audio signal (20 Hz-20 kHz).

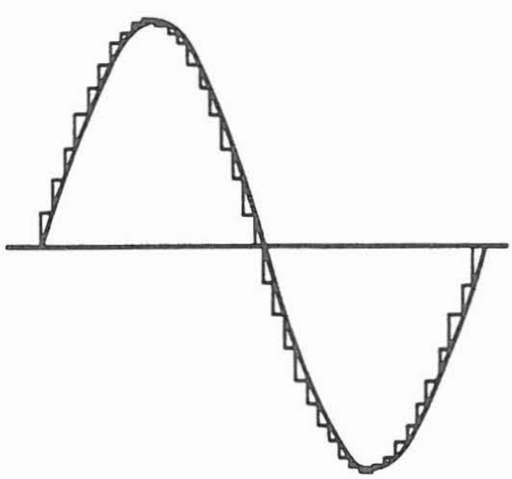
These high frequencies have to be suppressed by a low-pass filter (in the Compact Disc system their level has to be brought to at least 50 dB below the maximum audio signal level).
A suitable filter for this purpose is difficult to achieve in practice.

Digital filter
In order to allow mitigation of the low-pass filter requirements, the approach selected in the Philips Compact Disc player is to perform a preliminary filtering operation during the digital phase.
For this, a fourfold oversampling is performed, that is, a digital filter operating at four times the sampling frequency (4 x 44.1 kHz = 176.4 kHz) provides signal values at this raised frequency, thus refining the stepped curve and facilitating the filtering out of the high frequencies. As a result, it will suffice to use a relatively simple third-order low-pass filter after the digital-to-analogue conversion section.

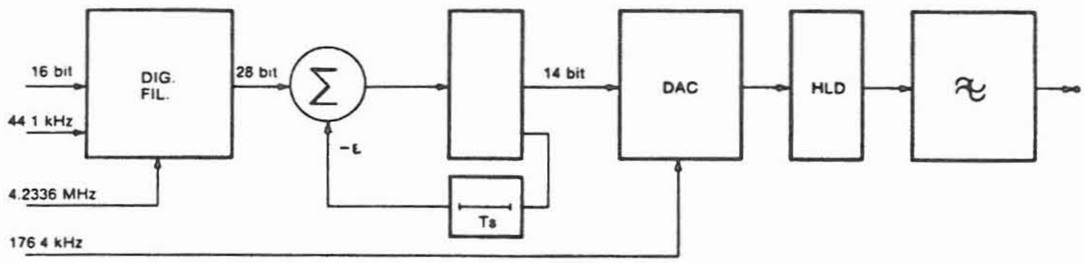
To convert the 16-bit words into an analogue signal the Philips Compact Disc player uses a 14-bit digital-to-analogue converter available in IC form; this DAC is designed to operate at the high sampling frequency of 176.4 kHz. The fourfold oversampling on the one hand and a feedback in phase opposition of the rounding errors on the other hand allows the rounding-off to 14-bits without degrading the signal-to-noise ratio (SNR). This SNR is maintained at about the 96 dB level associated with 16-bit quantization.



fs = 44.1 kHz



fs = 176.4 kHz



4.2336 MHz = 96 x 44.1 kHz

The set of operations belonging to the digital-to-analogue conversion are illustrated in the block diagram.
For stereo reproduction through two channels the complete chain is duplicated.
Oversampling takes place in the digital filter to which the input signal is fed. The next step consists in rounding off to 14-bits; the rounding error is fed back in phase opposition in the noise shaper (NS).
The digital filter and the noise shaper are accommodated in a single IC. Then follow the D/A converter and a hold circuit.
Finally, the analogue signal passes through a low-pass filter.

Digital-to-Analogue Converter (DAC)
The DAC is the device that reconverts the binary codes which represent the amplitudes of the audio samples into an analogue signal. Each binary code has its corresponding current.

Example
code 0000 provides a current of 0 mA, 0001 a current of 1 mA, 0010 2 mA, 0011, 3 mA, etc.

In the CD player the codes to be converted do not comprise four bits like in the example, but 14 bits; the number of different currents is 2^{14} = 16 384.

The 14-bit codes which enter the DAC control a number of switches which enable or inhibit the flow of these currents to the output.
In practice this will mean as increment of 0.6104 μA.

For a more detailed discription of the DAC used in the Philips Compact Disc see Appendix B.

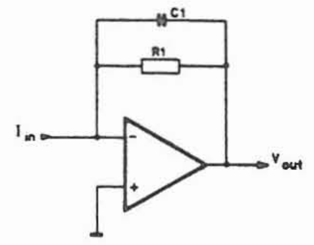
The Bessel filter

The application of the Bessel filter is shown in the Figure.
The filter for the second channel is identical with the one shown, except for the relay control circuitry which is required only once.
Bessel filters are normally designed by looking up the coefficients a and b of the normalized transfer function

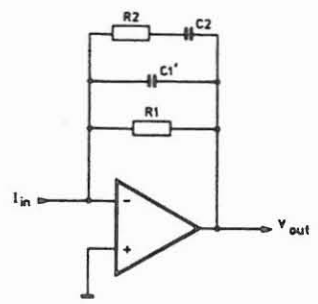
$$A(s) = \frac{1}{1 + S \, . \, a/\omega_C + S^2 \, . \, b/\omega_C{}^2}$$

in a table and then finding suitable component values.
ωc is the desired cut-off frequency of the complete Bessel filter which is equal to $2\pi \, . \, 30$ kHz.
The coefficients for a third-order Bessel filter are found to be $a_1 = 0.7560$ and $b_1 = 0.0000$ for the one-pole section and $a_2 = 0.9996$ and $b_2 = 0.4772$ for the two-pole section.

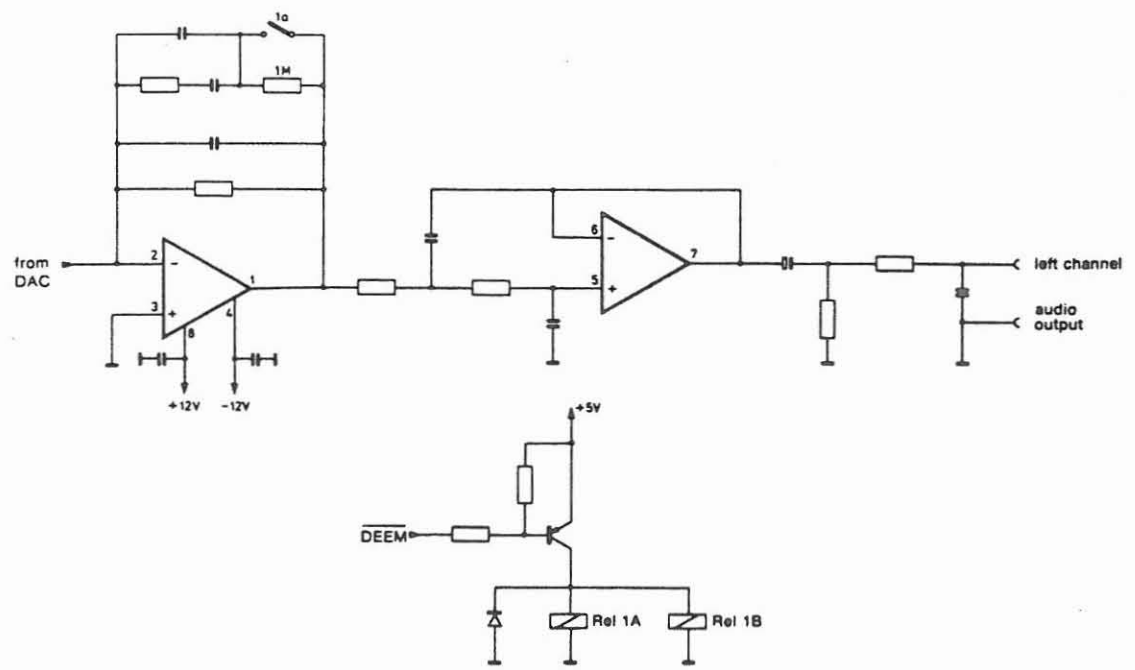
A de-emphasis characteristic is easily implemented in the one-pole section of the filter. This is achieved by closing the contact 1a.
If it is open the 1 MΩ resistor provides a discharge path for the capacitor but can be neglected for the calculation of the transfer characteristic of the filter. In the Figure the first part of the filter circuit is shown with and without a de-emphasis network.



without de-emphasis



with de-emphasis

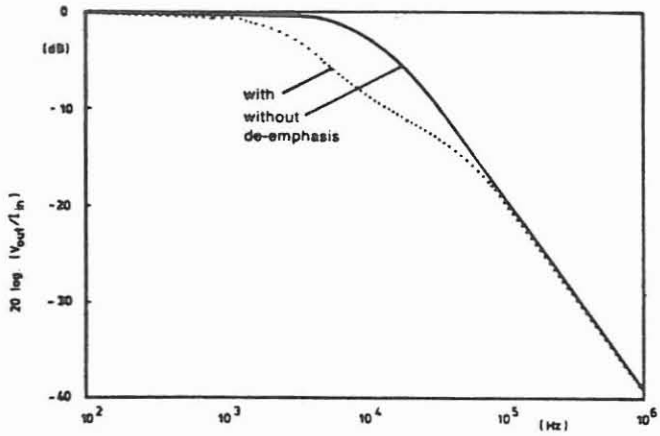


Besides converting an input current I_{in} from the DAC into an output voltage

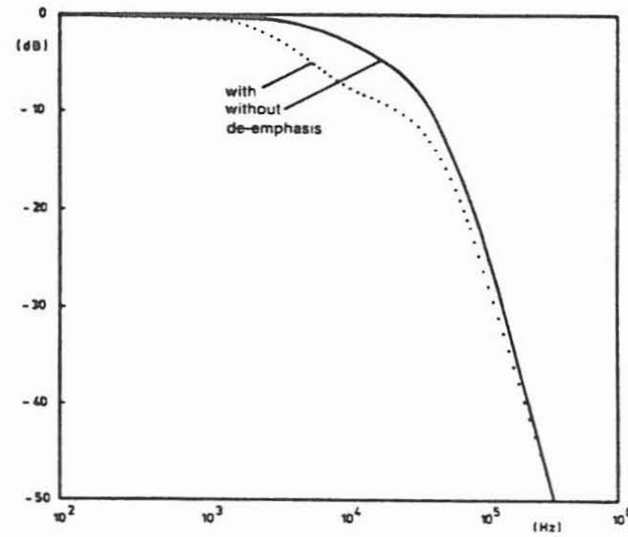
$V_{out} = -I_{in} \cdot R1$

the filter section provides a low-pass characteristic with a cutt-off frequency of about 40 kHz as shown in the curve. The first filter provides a first-order de-emphasis the frequency response of which is shown as a dotted line.

The frequency response of the complete Bessel filter is shown in the other curve. The lowest unwanted frequency in the output spectrum of DAC is 156.4 kHz which is 4 times the sampling frequency of 44.1 kHz minus 20 kHz. At this frequency the filter provides an attenuation of 33.8 dB. A further 18 dB are added due to the $\frac{\sin x}{x}$ effect of the D/A converter, so that the 176.4 kHz ± 20 kHz lobe is attenuated by more than 50 dB.



Frequency response of the first filter section



Frequency response of the Bessel filter

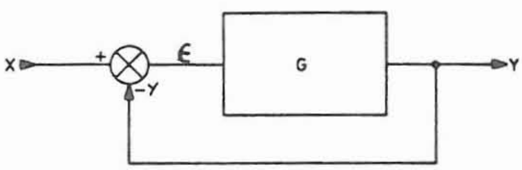
Servo systems

The principle of servo systems

Various mechanical movements in the CD player require control and verification. These movements are:

- radial tracking
- f··cusing
- .ntable motor speed

Control of these mechanical movements is performed by the various servo systems. To provide some insight into the problems which may arise in servo systems, we will start with an exposition of the principle of feedback systems.



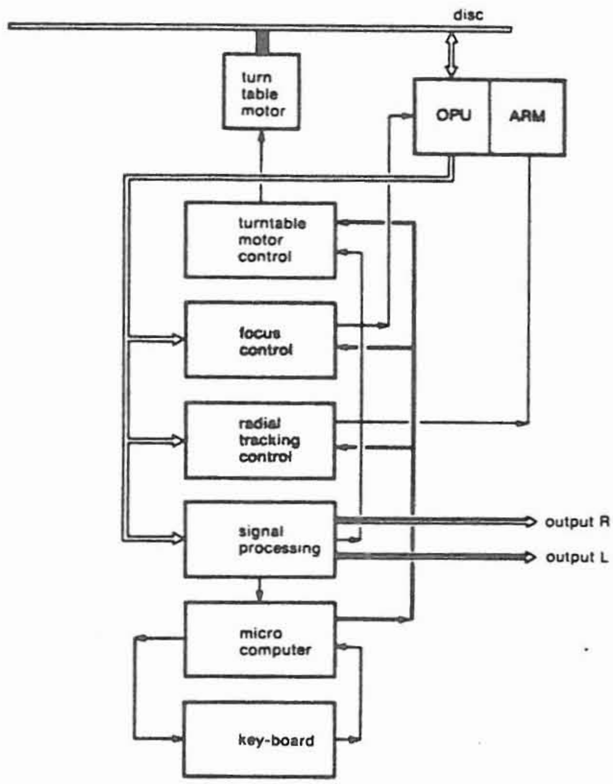
X = desired value (reference)
Y = actual value
ϵ = error signal
G = system properties

$$\left.\begin{array}{l}\epsilon = X - Y \\ Y = \epsilon \,.\, G \\ \epsilon = X - Y\end{array}\right\} Y = (X - Y)\,G$$

$$(I + G)\,Y = X.G$$

$$\frac{Y}{X} = \frac{G}{1 + G}$$

This formula leads to the conclusion that the system will become unstable when $G = -1$, because $\frac{Y}{X}$ will then be indeterminate.

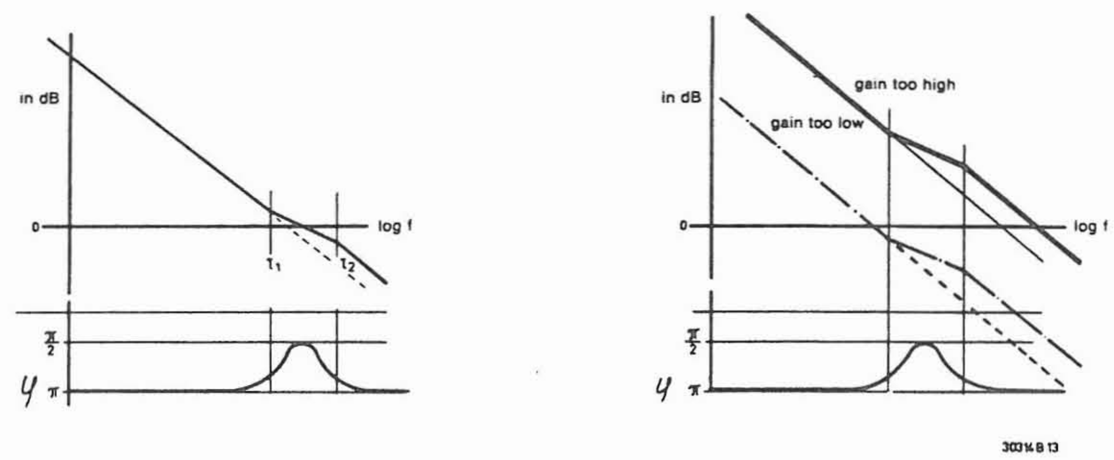


$G = -1$ is satisfied when $G = 1$, while at the same time a phase shift of 180 degrees is taking place.

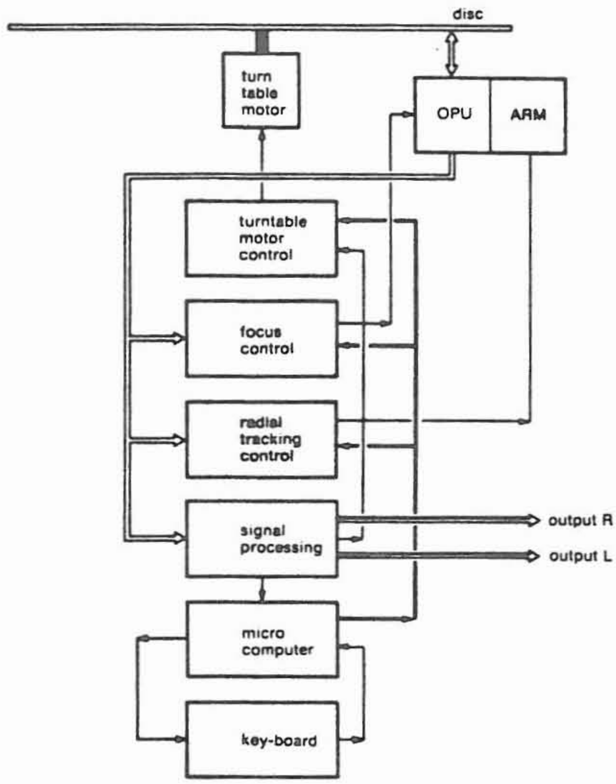
To prevent the system from becoming unstable, care must be taken that no 180° phase shift occurs when $G = 1$.

The transfer function of a system comprising a motor or something reacting like a motor, may be described as a 2nd order function with a roll-off of -12 dB per octave. However, a 2nd order system presents a phase shift of 180 degrees at $G = 1$. If no precautions are taken such a system will consequently become unstable each time the gain becomes 1. For this reason such systems are implemented with a so-called lead network that brings about a 90° phase shift with a roll-off of -6 dB/octave.
The time constants of the lead network must be chosen so, that the 90° phase-shift occurs when $G = 1$.
All the servo systems built into the CD player contain a lead network to keep them stable. The problem is, however, that the gain of the systems depends on so many factors that a change in these factors may result in finding the unity gain ($G = 1$) outside the crossover points of the lead network as shown in the Figure below.

Therefore it is advisable to keep the gain as constant as possible. This can be achieved by controlling the gain as is done for radial tracking (refer to following chapter).



The radial tracking system
The pick-up is attached to an arm that is capable of moving in a horizontal plane. The movement of the arm is based on the moving-coil principle as used in measuring instruments.
By passing more/less current through the coil, the arm can be brought to a defined position.



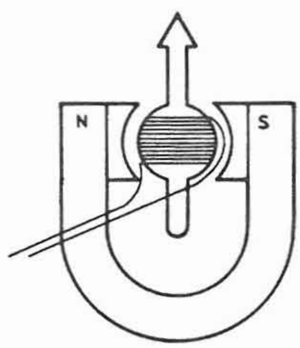
Information about the position of the unit relative to the track on the disc is obtained from the amount of light falling on the light-sensitive diodes.

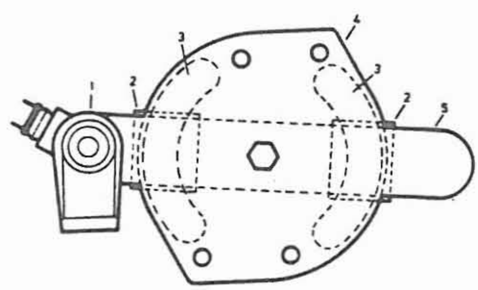
The diodes have been placed apart; in order to make the light fall on the diodes, the side of the prism where the reflected light emerges has been ground to a wedge shape. As a result of the wedge shape the reflected light is split into two equal beams. Each beam falling on a diode.
In reality we have diode pairs, but since for radial tracking the sum of the signals from each pair of diodes is used, we may as well speak of single diodes.

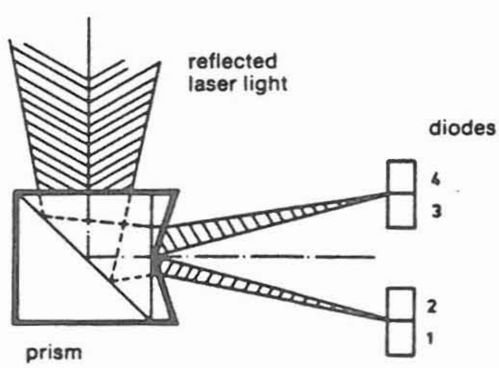
$(i_1 + i_2)$ is the signal from the first pair and $(i_3 + i_4)$ the signal from the second pair.

The radial error (RE) signal consists of the difference between the two signals:
$RE = (i_3 + i_4) - (i_1 + i_2)$.

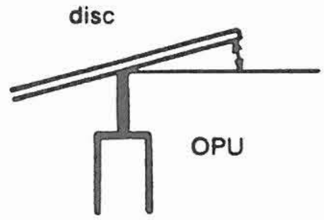
1 objective + lightpen
2 coil
3 magnet
4 chassis
5 arm
6 diodes



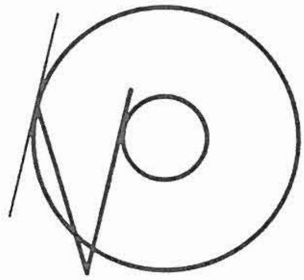




The signal may, however, be affected by a number of factors that could make the system unstable.



The factors are:
- a possible asymmetry of the spot;
- a possible deviation from the angle the disc must make relative to the pick-up unit;
- the arm's angle relative to the track (90 degrees in the centre of the disc and 45 degrees at the edge of the disc);
- the intensity of the spot;
- the reflectivity of the disc surface.



An asymmetric spot, resulting from a deviation in the angle between disc and pick-up unit, results in the signals $(i_1 + i_2)$ and $(i_3 + i_4)$ being unequal, even when the pick-up unit is below the centre of the track.
This error can be corrected by introducing a factor, called factor d (to be explained shortly).
As a result of the introduction of this factor d the radial error signal will become:

$$RE = 2d\,(i_1 + i_2 + i_3 + i_4) - 2\,(i_1 + i_2).$$

Compensation with factor d is not yet sufficient, for the system's gain is also affected by these factors.
If the laser produces little light, and the disc is reflecting well or badly etc., this may influence the gain by a factor of up to 6.
Such a variaton of gain can make the system unstable; to prevent this, the gain must be controlled.
The radial error signal will then become:

$$RE = k.d.\,(i_1 + i_2 + i_3 + i_4) - k\,(i_1 + i_2),$$

where factor k (also to be explained) and, consequent, the gain can be adjusted.

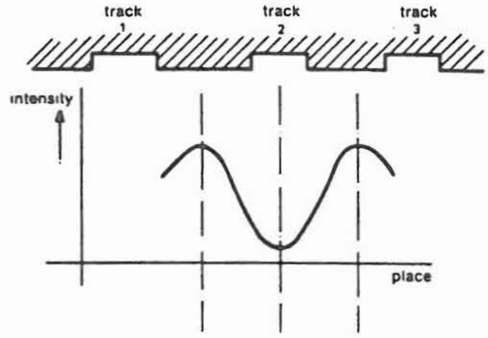
The questions that naturally arise are:

What are d & k, where do they come from, and how are they adjusted?

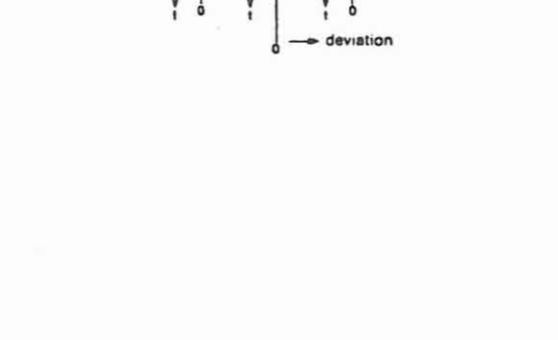
In the CD player use is made of an additional signal of 650 Hz which is injected into the radial system.
As a result of this signal the arm starts to make an oscillatory movement of 0.1 μm. Via the pick-up unit this signal is returned; it is then processed and passed to two multipliers.

Factor d
The curve as in the Figure opposite represents the amount of light reflected back from the disc as a function of the pick-up unit's place relative to the track.
This curve shows that much light is reflected when the unit is oriented between the tracks and little light when the unit directs the laser beam to the exact centre of the track. The effect to this curve on the 650 Hz signal that makes the arm oscillate, is shown in the Figure below.
When the laser light hits the centre of the track, the signal will be positive as a result of the 650 Hz signal.
A shift of the pick-up unit to the right of the track results in a returning signal that is in phase with the original 650 Hz signal.
If the laser light lands at the left of the track centre, the returning signal is in phase opposition with the original 650 Hz.

By detecting synchronously the returning signal, a control signal is obtained that determines the factor d.
The factor d co-determines the radial error signal that in turn sets the arm into movement so that the laser spot is correctly directed to the centre of the track.



Factor k
For the adjustment of factor k use is made once again of the 650 Hz signal.
It appears that in the radial tracking system the phase of the 650 Hz signal is strongly influenced by the gain.
If the gain increases or decreases the phase shift between the original signal and the returning signal will also increase or decrease.
Both signals are compared with each other in a phase detector and the resultant signal is proportional to the phase shift and, consequently, to the magnitude of the gain.
This resultant signal is used to adjust factor k, thus maintaining the gain of the entire system as constant as possible.



Offset compensation radial tracking

om the explanation of the radial tracking followed that the radial error signal is found from

$RE = (i_3 + i_4) - (i_1 + i_2)$.

However if the light spot is not symmetrical or the disc-light beam angle is not 90 degrees, this signal will not equal zero when the spot is in the track centre, because $(i_3 + i_4)$ will not be equal to $(i_1 + i_2)$.

The difference between the signals $(i_3 + i_4)$ and $(i_1 + i_2)$ can be compensated for by multiplying $(i_1 + i_2)$ by $(1 + E)$ and dividing $(i_3 + i_4)$ by $(1 + E)$ resulting in:

$$RE = \frac{(i_3 + i_4)}{(1 + E)} - (i_1 + i_2)\,(1 + E)$$

As the difference will be small: $|E| \ll 1$

With some mathematics can be found:

$$RE = \frac{(i_3 + i_4)\,(1 - E)}{(1 + E)\,(1 - E)} - (i_1 + i_2)\,(1 + E) = \frac{(i_3 + i_4)\,(1 - E)}{1 - E^2} - (i_1 + i_2)\,(1 + E)$$

$E \ll 1$ so $RE = (i_3 + i_4)\,(1 - E) - (i_1 + i_2)\,(1 + E)$
(E^2 can be neglected)

Suppose: $1 - E = 2d$; thus $E = 1 - 2d$ and $1 + E = 2\,(1 - d)$

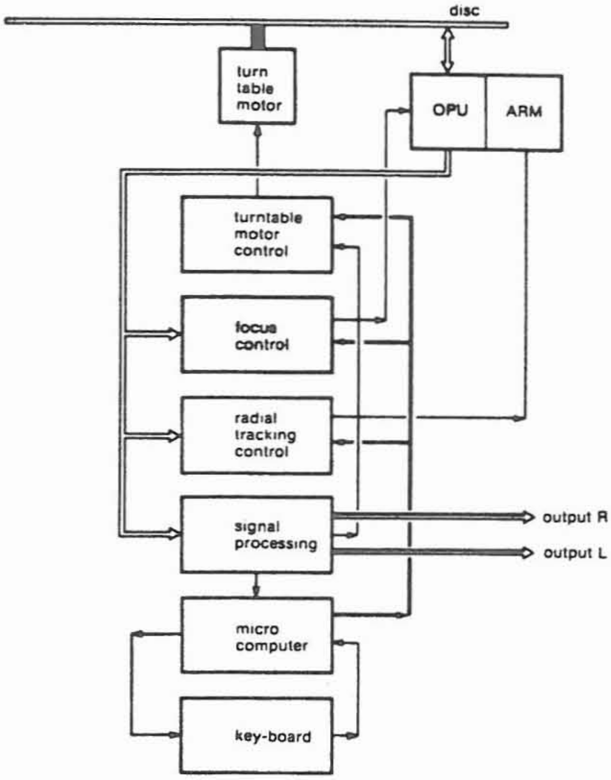
'han is:

$$RE = 2d\,(i_3 + i_4) + 2d\,(i_1 + i_2) - 2\,(i_1 + i_2)$$

$$RE = 2d\,(i_1 + i_2 + i_3 + i_4) - 2\,(i_1 + i_2)$$

The focus servo system

In the description of the general block diagram it was said that a position of the disc above or below the objective's focal point will make the focal point of the reflected light land behind or in front of the diodes.



The shape of the prism causes the reflected light to be split up into two equal beams, each beam falling on a diode pair.
If the system is correctly focused, that is, the distance between objective and reflective surface is nominal, the spot lands at the centre between the two diodes of the diode pair.

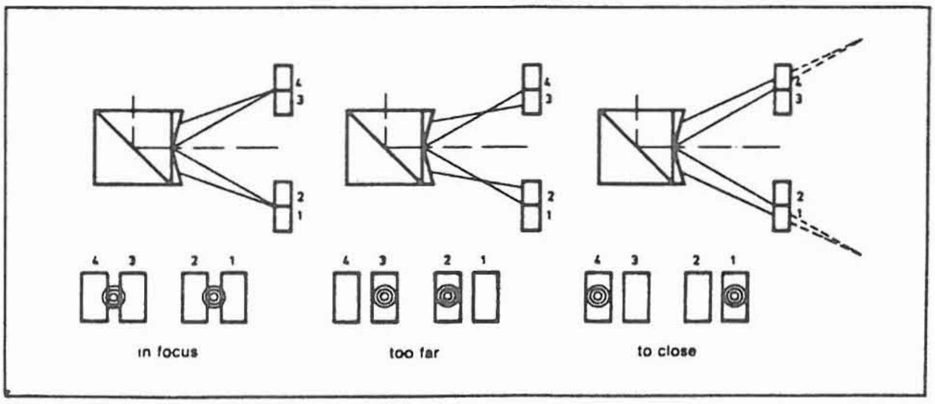
In this case both diodes receive an equal amount of light. If, on the contrary, the distance between objective and disc is too great, so that the focal point of the reflected light moves to a point in front of the diodes, the light spots will land on the diodes 2 and 3. In the opposite case, that is, when the distance is too little, the focal points will be somewhere behind the diodes, and the diodes 1 and 4 will receive the greatest quantity of light. By summing the signals from the diodes 2 and 3 and those from the diodes 1 and 4, and by subtracting the results, the focus error signal is found.

$FE = (i_1 + i_4) - (i_2 + i_3)$.

From the outcome of this formula the following conclusions can be drawn:

$FE = 0$: the system is correctly focused.
$FE < 0$: distance between disc and objective too great
$FE > 0$: distance between disc and objective too small

The focus error signal (FE) is, via an amplifier, fed to the coil that causes the objective to move up or down until the FE signal is again 0 and the system correctly focused.



Summary

The Compact Disc player reads out the information contained in the disc and distils from it the audio signal which can be reproduced via an amplifier and a few speakers.
The information is present on the disc in the form of pits or depressions in a reflective layer. With the aid of a laser the information is read out.
The laser beam is focused on the reflective layer that reflects back all of the beam. If the laser beam falls on a pit, the intensity of the reflected light will be low. In this way the reflected beam is intensity-modulated by the information on the disc.
This light is via a prism projected on light-sensitive diodes which in turn convert the intensity variations into electric signals.
The signals from the diodes are used for various purposes i.e.:
- to be converted into audio signals;
- for focusing;
- for track following.

Conversion of the diode signals into audio signals
The diode signal (HF signal) enters the demodulator IC.
A PLL circuit allows to process this signal to have a sync signal for information readout at the correct moments of time.
Using the sync pattern in the HF signal, the readout is synchronized and the digital codes can be read. Finally these codes are transformed from 14-bit codes to 8-bit codes and passed to the error correction IC.

The first check performed by the error correction IC is that of the speed at which the DATA enters. For this purpose the speed of the incoming data flow is compared with a very stable clock frequency. The comparison results in the turntable motor control signal.

Next, the 8-bit code words which contain the audio information, are stored in a memory (RAM) before being passed to the error correction decoders. This memory makes it possible to present the codes or symbols to the decoder in the order necessary to accomplish the decoding and error correction process.
In the first decoder the 32 symbols of a frame are split up into 28 information symbols and 4 parity symbols.
The following step is to multiply the parity symbols matrixwise by the information symbols. If no errors are found, the 28 symbols return to the memory.
If not more than one error is found, this error is corrected and the corrected frame is returned to the memory.
If more than one error occurs, the 28 symbols are placed back uncorrected in the memory; however, error signals (flags) are generated which inform the second decoder that these symbols are wrong. Since in this situation a de-interleaving action must be undertaken, the symbols arrive in a different order in the second decoder.
Four out of the 28 symbols presented are parity symbols, the remaining 24 contain audio information.
Again matrixwise multiplications are carried out.
If no errors are found, the 24 symbols go back to the memory. If wrong symbols occur (up to a maximum of two), these are corrected and the corrected frame is then placed back in the memory.
In the event of more than two wrong symbols, the frame is placed back unchanged in the memory; at the same time error signals are generated which instruct the next IC which symbols are wrong.
The next IC is the interpolating and muting IC.
With the aid of the flag from the error

correction IC the interpolating and muting IC is capable of determining whether an erroneous sample can be interpolated or ould be muted.

In cases where the values of the preceding and following samples are known, linear interpolation makes it possible to compute a value for the missing sample. If one of these sample values is absent, muting is applied.
A further task of the interpolating and muting IC is to separate the information for the right and left channels so that two outputs are available.
Each channel signal is then fed to a digital filter that sees to it that the audio frequencies only are let through to the digital-to-analogue converter (DAC).
The DAC generates an analogue current which corresponds with the binary word at the input. This current changes each time a different binary code appears at the input. In s way the original analogue audio signal is constructed from the binary codes.
ally, an analogue filter is used to further suppress any interfering frequencies.

Radial tracking
When the laser beam is incorrectly centered on the track, the intensity of one half of the reflected beam will be lower than that of the other beam. The split up of the reflected beam into two equal parts and the falling of each half on a separate diode makes it possible to detect whether the track is followed correctly.
If an equal amount of light falls on either diode, we know that the track is correctly followed. If the light beam is incorrectly centered on the information track, the light one diode will be less than on the other ode.
taking up the arm that bears the pick-up unit into a servo system it is possible to control the arm in a way that maintains equal amounts of light falling on each diode.
The spot intensities are also affected by factors such as: asymmetry of the spot, deviation from the ideal angle between disc and pick-up unit.
The introduction of factor d allows compensation for these influences.
To ensure the stability of the servo system, it is also necessary to control the loop again. For this purpose the radial error signal is multiplied by a factor k which is dependent on the gain.

Expressed as a formula, the radial error signal will then be:

$$RE = k.d\,(i_1 + i_2 + i_3 + i_4) - k\,(i_1 + i_2).$$

The factors d and k are obtained by means of a 650 Hz signal that is injected into the servo system.

Provision for good focusing
When the laser beam is correctly focused, the beam's focal point coincides with the reflective surface.
In that case the focal point of the reflected light coincides precisely with the plane of the diodes.
The diodes are arranged so that each deviation from correct focusing is translated into a focus error signal. This focus error signal is used to control the objective so that correct focusing is restored and the focus error signal becomes again zero.

Appendix A

Compact Disc encoding

Questions

System basics

1. *Why is a sharp cut-off low-pass filter required?*
2. *What happens in the Sample-and-Hold circuit?*
3. *How is quantization noise created?*
4. *What takes place in the ADC?*
5. *Where are the necessary actions for error correction performed?*
6. *What is a synchronization pattern and what is it needed for?*

Error detection and correction principles

1. *What are redundancies and what are they needed for?*
2. *In how many positions must valid code words differ to enable detection of t_1 errors?*
3. *What is meant by the minimum distance (d_m)?*
4. *In how many positions must valid code words differ to enable correction of t_2 errors?*
5. *What formulas apply for error detection and correction of symbols in frames?*
6. *How many parity symbols are added to the audio symbols by the encoding system?*

Scrambling and interleaving

1. *What is the result of interleaving with regard to the symbols on the disc?*
2. *How is interleaving brought about?*
3. *What is the effect of interleaving on the error encoding system?*
4. *What is meant by interpolation in the Compact Disc field?*
5. *What is meant by scrambling?*
6. *What are the consequences of scrambling on the locations of the symbols on the disc?*
7. *How does the frame look, the way it is recorded on the disc?*

Modulation

1. *What requirements must the modulation system for the Compact Disc fulfill?*
2. *In what way is regeneration of the clock frequency achieved mostly?*
3. *What are the consequences of the light spot's finite dimensions?*
4. *Which encoding procedure, used in the Compact Disc, meets the modulation system's requirements?*
5. *Give two reasons for introducing merging bits?*

Compact Disc encoding

Introduction

This Appendix is dealing with the way in which the analogue audio signals are processed prior to recording on the disc. Ample attention will be given to the principles of error correction so as to enhance the understanding of the error correction IC.

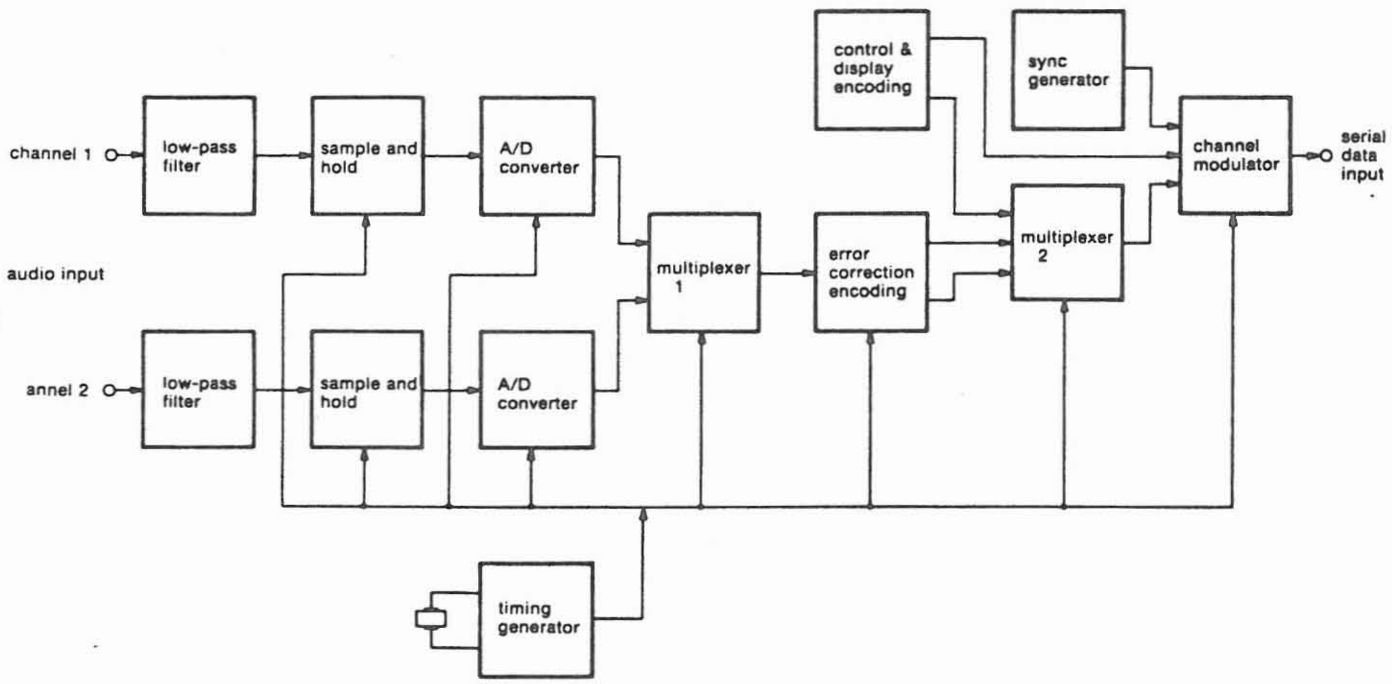
The Compact Disc (CD) is a storage medium for the reproduction of highest-quality audio signals. Sound is digitized and recorded on a disc as groups of binary numbers represented as pits and flat areas along a helical track.

The change from the analogue to the digital domain for the processing of audio signals leaves all analogue problems like wow and flutter behind. There are, however, two new problems related to digitized audio: quantization noise, i.e. errors inherent in the conversion of analogue signals to digital, and drop-outs, i.e. pulses lost due to transmission failures.

Quantization noise can be reduced to a negligible level by increasing the word length of the digitized audio signals to a sufficient length, 16-bits are used in the CD system. The occurence of drop-outs is a much more severe problem. Intensive investigations have been carried out in the field of error detection and error correction. As the required circuitry is very complex, a set of LSI circuits has been developed for the cost-effective realisation of a decoder for the Compact Disc Digital Audio system.

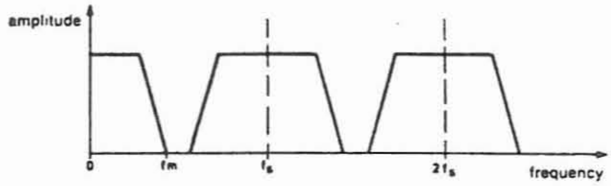
System basics

The CD carries digitized audio information for a playing time of up to one hour in stereo. For the production of the disc, the audio signals have to be digitized and encoded. The block diagram of this encoding system is shown in the Figure.

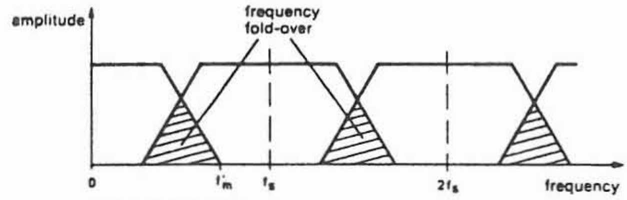


The low-pass filters

Before any other processing takes place, the analogue input signals are passed through a sharp cut-off low-pass filter in order to limit their bandwidth to a maximum frequency f_m equal to or less than one-half the sampling frequency of $f_s = 44.1$ kHz. This is to ensure conformity with the Nyquist theorem.
Otherwise, intermodulation distortion could occur due to frequency fold-over.
In an impulse-sampled system, the frequency spectrum simply is the audio signal spectrum repeated periodically with a scaling factor of f_s.



with band-limiting

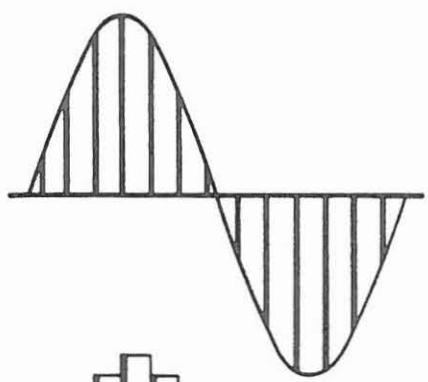


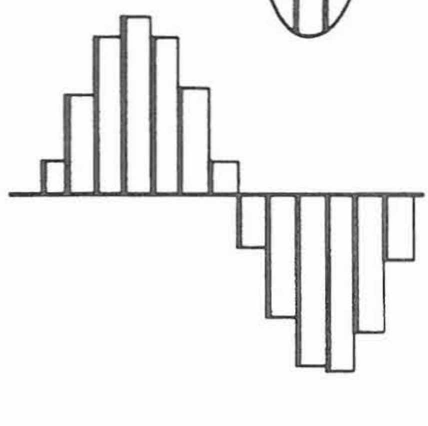
without band-limiting

Sample-And-Hold

Before the stereo audio signal is recorded on the disc, it has to be converted into a digital signal.

One of the first conversion steps is to sample the analogue audio signal at fixed intervals or time points.
At each of these time points (also referred to as sampling times) the signal's amplitude is measured and the measured value is held for a moment to enable its processing, like - for instance - the conversion of the measured value into a binary code.
This way of sampling is referred to as Sample-And-Hold technique.

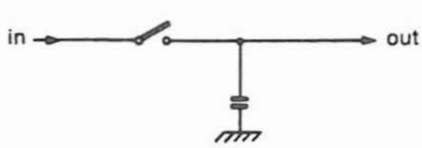




The procedure of this Sample-And-Hold technique is very much similar to the charge/ discharge procedure of a capacitor via a switch.

The switch is actuated by a pulse of a relatively short duration. This pulse is named 'sampling pulse'.

When the analogue signal comes in at the INput and the switch is closed by the sampling pulse, the capacitor will be charged to the signal's amplitude at the moment (= instantaneous value).
If the switch is then re-opened, the capacitor will retain the instantaneous value until the switch is closed again. In this new situation the capacitor will take over the signal's present instantaneous value.



The maximum frequency that can be sampled in this way is one-half the frequency of sampling.
In the CD system the sampling frequency is 44.1 kHz and, consequently, amply sufficient for a signal frequency of 20 kHz.

Analogue-to-digital conversion

After sampling of the analogue signal, the amplitude value has to be converted into a binary code.
The problem is, however, that the analogue signal may assume an infinite number of levels, whereas the number of binary codes available to reproduce the analogue levels is finite.
The sampled signal must therefore be quantized, that is, the maximum amplitude which may occur is divided into a number of levels equal to the available number of binary codes.
The real value of the analogue signal has now to be rounded to a quantized value that comes closest to the analogue value .
The difference between the signal's real value and the quantized signal is experienced as noise and is referred to as quantization noise.
The greater the number of codes available for the purpose, the greater the number of steps and the smaller their interval size, thus leading to a reduction of the quantization noise.
The Figure shows how quantization noise is created.
The error signal is the difference between the input signal and the quantized signal curve.
Maximum error is ± ½ LSB (Least Significant Bit).
For the computation of the SNR the quantization noise is commonly assumed to be identical to white noise.

In formulated form the signal-to-noise ratio (SNR) is:

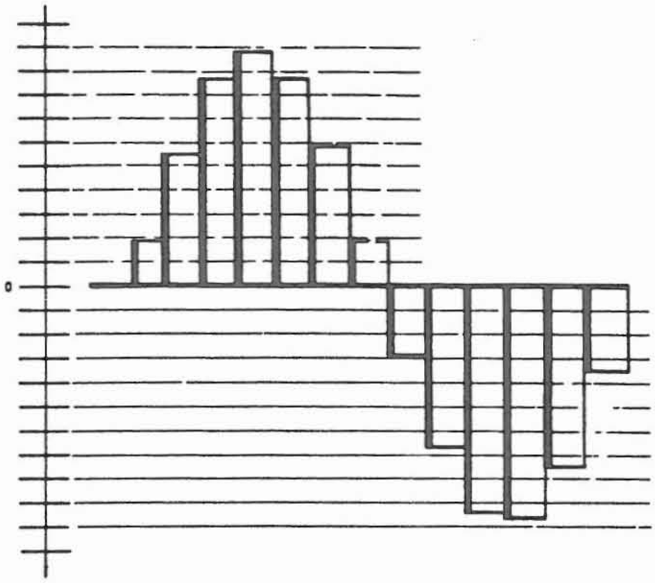
$$SNR = 20 \log N + 1.76 \text{ dB}$$

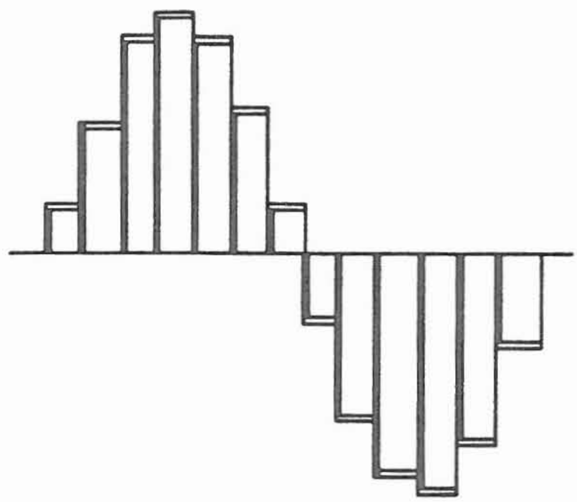
where N = the number of available codes.

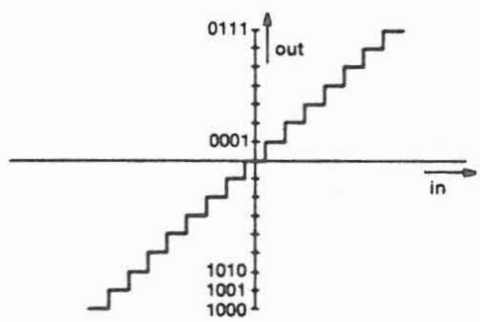
For the Compact Disc system: $N = 2^{16}$, since a code consists of 16 binary digits or 16 bits.

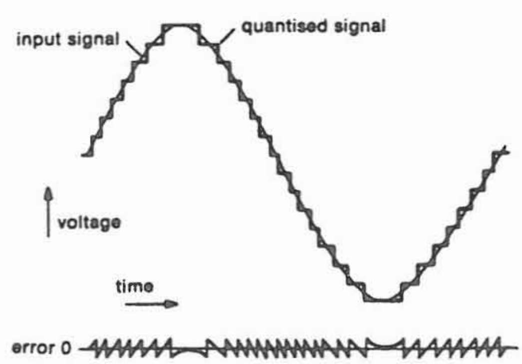
Thus dividing the amplitude into 65536 equal, intervals gives
$SNR = 20 \log 2^{16} + 1.76 = 98$ dB.

Each quantized level corresponds to a binary code.
The conversion of the quantized level into a binary code is carried out in an analogue-to-digital converter (ADC).



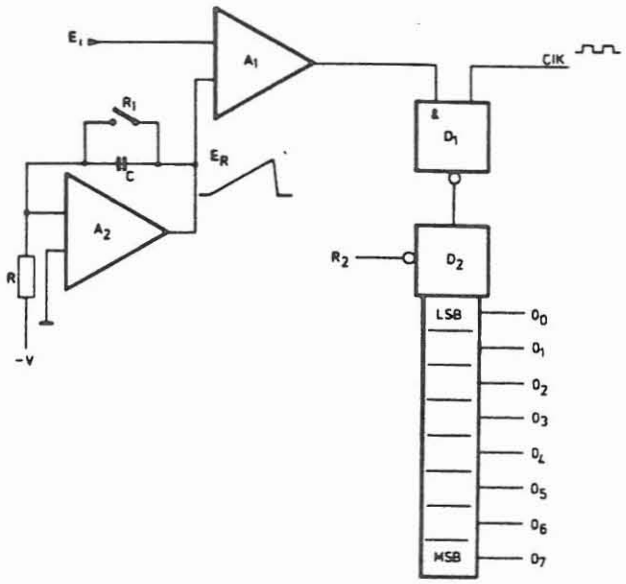




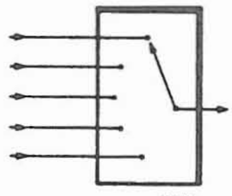


The principle of operation of an ADC is as following:

One-slope ADC
The analogue input voltage E_i is presented to comparator A1.
A1 compares E_i with E_r (E_r is the sawtooth voltage coming from integrator (A2).
When E_i and E_r are equal, the output of A1 switches over and blocks the clock pulses Clk via NAND gate D1.
The binary counter D2 stops and he binary content of this counter represents the digital conversion of an analogue value.



Multiplexer
Sampling, quantization and conversion to a binary code are operations performed separately for the right and left channels. However, both codes should ultimately be stored on the disc, one after the other in a single track.
For this, use is made of a multiplexer; the multiplexer is a digital circuit that first passes on the information of one channel and then the information of the second channel, etc.



Multiplexer 1
The 16-bit binary words coming out the A/D converters are fed to a multiplexer which in turn passes them sequentially to the error correction encoding unit.

Error correction encoding
As any errors in the transmission of digitized audio signals would result in audible effects, it is of paramount importance to detect and correct errors as far as possible. The method used with the CD system is based upon parity bits and interleaving. It is termed Cross Interleaved Reed-Solomon Code (CIRC).
The necessary actions are performed by the error correction encoding unit. There are two outputs from this unit, one carries data and the other parity symbols.

Control and Display encoding
As the CD offers a new dimension in sound quality, it seems to be justified to equip a CD player with operating features which cannot easily be found in conventional record players. This will further aid in the design of CD players with an outstanding performance.

An identification of the pause between any two successive pieces of music can be used for the implementation of search and repeat functions, an identification whether the recording was made with pre-emphasis or not could be used for switching the de-emphasis in the player on and off automatically, recording of timing information could be used for the display of elapsed time or of the time of play or a piece of music, to mention just a few potential features.
As this control and dispay information is non-audible, it has to be encoded separately. This is performed by the control and display encoding unit which outputs 8-bit wide symbols, permitting the implementation of eight different information channels two of which have been specified so far.
The control and display information is sometimes referred to as subcoding information.

As the subcoding information is generated and recorded bit-serially (seen per channel), there must be some provision for the recognition of the beginning of a string of bits.
There are two outputs from this encoding unit, one carries data and the other is used for the synchronization of the subcoding data blocks.

Control and display system

After demodulation 8 bits per frame (P, Q, ...W) are available for control and display purposes.

These 8 bits are used as 8 different subcoding channels, so giving each channel a bitrate of 7,35 kBit/sec.

Channel P is a simple music piece separator flag. Cheap search systems can be implemented by only using this flag.

Channel Q is used for control purposes of more sophisticated players. The decoding of this channel can be implemented with a μ-computer. In channel Q are encoded items lide track number and time.

The minimum length of a piece of music is 4 seconds (not including the pause length before the piece of music).

The Figure gives an example of the encoding in channel P and Q.

The channels R..W can be used for display purposes.

Date format

```
subcoding symbol      PQRSTUVW
                       d1 ...... d8
example                00101111       NRZ notation
                           |
                           v
after modulation     |C1 ------------ C14|
                     | 00100001001000 | XXX | (NRZ-I notation)
                           |
                           |            mer-
                           |            ging
                           v            bits
          sync       | control & display     | symbol-1

                     ------------------------> time
                     bit C1 is first out
```

The frame structure

he subcoding block consists of 98 suboding symbols. The repetition frequency of one block is 75 Hz.
The first two subcoding symbols are the subcoding sync patterns S0 and S1.

d1 ------------------------ d8

	P	Q	R	S	T	R	V	W
0			sync pattern S0					
1			sync pattern S1					
2								
.								
.								
.								
97								

S0 = (00100000000001)
   = (00000000010010)
   C1 ... C14

Channel P

Channel P is a flag bit that indicates the start of a music piece with the following code rules:

music : P = 0
start flag : P = 1

The minimum length of the encoded start flag in channel P is 2 seconds, the end of the encoded start flag gives the start of the next music piece.
he actual pause exceeds 2 seconds, the ngth of the start flag gives the actual pause ngth.

Lead-in track

In the lead-in track channel P is encoded as music. The first piece of music is preceded by a start flag of 2-3 seconds.

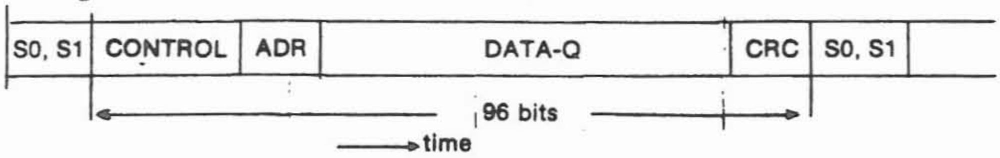
Lead-out track

The lead-out track is preceded by a start flag of 2-3 seconds (during the last music piece on the disc). The end of the start flag gives the begin of the lead-out track. Channel P remains zero for 2-3 seconds after the start of the lead-out track, next P switches between 0 and 1 in a 2 Hz ± 2% rythm (duty cycle 50 ± 10%).

hannel P can only change after the subcoding sync pattern S0, S1.
The encoding of channel P is one subcoding block delayed to the encoding of channel Q.

Channel Q

The general data format of channel Q is:



CONTROL: 4 flag bits for the number of channels and pre-emphasis on/off, MSB is first out.

- 0000 - 2 audio channels without pre-emphasis.
- **1000 - 4 audio channels without pre-emphasis.**
- **x001 - 2 or 4 audio channels with pre-emphasis of 50/15 μs.**

The pre-emphasis flag (bit 4 of the control block) can only change during an actual pause (X = 00) of at least 2 seconds. During the lead-in and lead-out tracks, the pre-emphasis flag is zero.

ADR 4 control bits for DATA-Q, MSB is first out.

DATA-Q 72 bits DATA, always MSB first out.

CRC A 16 bit CRC on CONTROL, ADR and DATA-Q, MSB first out. On the disc the parity bits are inverted. The remainder have to be checked at zero.
Polynomial:

$$P(X) = X^{16} + X^{12} + X^{5} + 1$$

Three modes are defined for DATA-Q.

Mode-1

ADR = 1 = (0001)
Mode-1 occupies at least 9 out of 10 successive subcoding blocks.
Two different data formats are possible in mode-1.

During the lead-in track, the data format is:

S0,S1	CONTROL	1	00	POINT	MIN	SEC	FRAME	ZERO	PMIN	PSEC	PFRAME	CRC
		ADR	MNR									

During the music and lead-out tracks on a dish, the data format is:

S0, S1	CONTROL	1	MNR	X	MIN	SEC	FRAME	ZERO	AMIN	ASEC	AFRAME	CRC
		ADR										

NMR Music number expressed in 2 digits BCD.

00: Lead-in track.
The end of the lead-in track is at the starting diameter of the program area.

01-99: Music numbers.
A music piece can be preceded by a pause with the same music number. The music numbering has to start with the value 01 and has to increment by one. In case a program is stored on several discs, the numbering may be continued.

AA: Lead-out track.
The lead-out track starts at the end of the last music piece on a disc, without a preceding pause encoding.

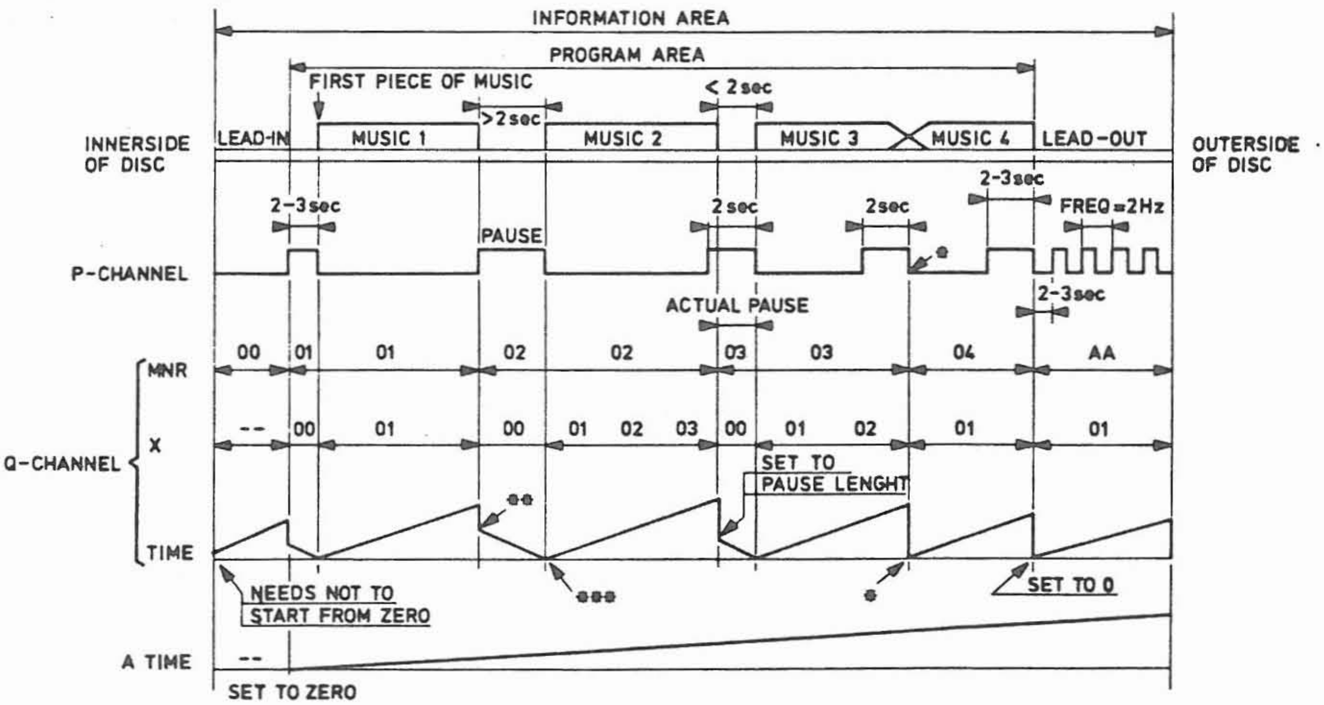
X	Index to MNR, 2 digits BCD. During the lead-in track, the index X is not available. 00: Pause encoding. The pause encoding in channel Q gives the pause length like it is in the music. The first piece of music is preceded by a pause encoding of 2-3 sec (see channel P). The lead-out track is encoded as music. 01-99: Subdivision numbers. During the lead-out track X is 01. Within a music piece (MNR = 01-99 and X = 00) the first value of X is 01. The value of X only can increment by one.
ZERO	These 8 bits are zero.
MIN SEC FRAME	Running time within a piece of music expressed in 6 digits BCD. MIN, SEC and FRAME each 2 digits. The time is set to zero at the start of a music piece. Time increases in the music and decreases in the pause, ending with the value zero at the end of the pause. In the lead-in and the lead-out tracks the time increases. The minutes are stored in MIN, the seconds in SEC. One second is subdivided into 75 FRAME (running from 00 to 74).
AMIN ASEC AFRAME	Running time on the disc expressed in 6 digits BCD. AMIN, ASEC and AFRAME each 2 digits. At the starting diameter of the program area the running time is set to zero and MNR takes the value of the first music piece on the disc. The minutes are stored in AMIN the seconds in ASEC. One second is subdivided into 75 AFRAME (running from 00 to 74).
POINT PMIN PSEC PFRAME	During the lead-in track a table of contents is stored into these locations. This table of contents is continuously repeated in the lead-in area (MNR = 00). In each table of contents, the items are repeated three times. At the end of the lead-in area, the table of contents can be ended with any value of POINT. The value of PMIN, PSEC and PFRAME gives the starting point of the music number pointed to by POINT. These values give the start position of the music piece on the absolute time scale (AMIN,

ASEC and AFRAME) with an accuracy of +/— one second. The start position of a music piece is the first position with the new music number and X = 00.
If POINT = A0, the value of PMIN gives the MNR of the first music piece on the disc, PSEC and PFRAME are zero.
If POINT = A1, the value of PMIN gives the MNR of the last music piece on the disc, PSEC and PFRAME are zero.
If POINT = A2, in PMIN, PSEC and PFRAME the starting point of the lead-out track is given.

The Figure gives an example of the encoding of the table of contents with 6 music pieces on a disc.

Frame number	POINT	PMIN, PSEC, PFRAME
n	01	00,02,32
n+1	01	00,02,32
n+2	01	00,02,32
n+3	02	10,15,12
n+4	02	10,15,12
n+5	02	10,15,12
n+6	03	16,28,63
n+7	03	16,28,63
n+8	03	16,28,63
n+9	04	. .
n+10	04	. .
n+11	04	. .
n+12	05	. .
n+13	05	. .
n+14	05	. .
n+15	06	49,10,03
n+16	06	49,10,03
n+17	06	49,10,03
n+18	A0	01,00,00
n+19	A0	01,00,00
n+20	A0	01,00,00
n+21	A1	06,00,00
n+22	A1	06,00,00
n+23	A1	06,00,00
n+24	A2	52,48,41
n+25	A2	52,48,41
n+26	A2	52,48,41
n+27	01	00,02,32
n+28	01	00,02,32
.	.	. .
.	.	. .

EXAMPLE OF ENCODING IN CHANNEL P AND Q



* This point can be determined by the software maker.

** The stop point of the music in the player is the location where the MNR changes.

*** The start point of the music in the player is the first location with the new MNR and the X = 00.

The accuracy of the start and stop points depends on the player design.

The switch delay of the preemphasis depends on the player design.

— — During the lead-in track. ATIME and X are not available.

Mode-2

ADR = 2 = (0010).
If mode-2 is present, it occupies at least 1 out of 100 successive subcoding blocks.

The data format in mode-2 is:

S0,S1	CONTROL	2	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	N11	N12	N13	ZERO	AFRAME	CRC
		ADR																
			← 52 bits →															

N1-N13: Catalog number of the disc expressed in 13 digits BCD.
Used is the UPC/EAN-code (BAR coding).
The catalog number does not change on a disc. In case no catalog number is encoded according to the UPC/EAN-code, N1-N13 are all zero, or mode-2 can be deleted from the disc.

ZERO: These 12 bits are zero.

AFRAME: The continuation of AFRAME in mode-1 (two digits BCD running from 00 to 74).
During the lead-in area (MNR = 00), these 8 bits are zero.

Mode-3

AL. = 3 = (0011)
If mode-3 is present, it occupies at least 1 out of 100 successive subcoding blocks.
Mode-3 is used to give a unique number to a piece of music. This is done by means of the International-Standard-Recording-Code (ISRC). The ISRC is defined in DIN-31-621).
If no ISR Code is used, mode-3 must be deleted.
During the lead-in and lead-out tracks, mode-3 is not present on the disc.
The ISR Code only can change immediate after the MNR has been changed.

The data format in mode-3 is:

S0,S1	CONTROL	3	I1	I2	I3	I4	I5	00	I6	I7	I8	I9	I10	I11	I12	ZERO	AFRAME	CRC
		ADR	ISRCode 60 bits															

T. 12 characters of the ISRC code are represented by I_1-I_{12}.
The Country-code is given in I_1-I_2,
I_3-I_5 give the Owner-code, I_6-I_7 the year of recording and I_8-I_{12} the serial number of the recording. The characters I_1-I_5 coded in a 6-bit format according to the figure. The digits I_6-I_{12} are 4 bit BCD numbers.

Character	*Binary*	*Octal*	*Character*	*Binary*	*Octal*
0	000000	00	I	011001	31
1	000001	01	J	011010	32
2	000010	02	K	011011	33
3	000011	03	L	011100	34
4	000100	04	M	011101	35
5	000101	05	N	011110	36
6	000110	06	O	011111	37
7	000111	07	P	100000	40
8	001000	10	Q	100001	41
9	001001	11	R	100010	42
A	010001	21	S	100011	43
B	010010	22	T	100100	44
C	010011	23	U	100101	45
D	010100	24	V	100110	46
E	010101	25	W	100111	47
F	010110	26	X	101000	50
G	010111	27	Y	101001	51
H	011000	30	Z	101010	52

00 These 2 bits are zero.

ZERO: These 4 bits are zero.

AFRAME: The continuation of AFRAME in mode-1 (two digits BCD running from 00 to 74).

Channels P, S, T, U, V, W

These channels are zero, it is not permitted to write other data into these locations.

Multiplexer 2
The 8-bit symbols from the error correction encoding unit which are either data or parity symbols and from the control and display encoding unit are passed through a second multiplexer which in turn feeds them in a certain sequence to the channel modulator.

Sync generation
As the data are organized in blocks, some means for recognizing the beginning of these blocks must be provided. The sync generator creates a unique pattern which is not contained in the normal data. It is passed to the channel modulator which adds it to its output data on request from the timing unit.

Channel modulator
The output signals of the error correction encoding and control and display encoding units are both represented in a non-return-to-zero (NRZ) format.
These NRZ data may not be recorded on the disc.
They have to be changed to another code, an action which is performed by the channel modulator.

Timing
As we have seen, many different tasks have to be carried out for the encoding of a CD. These tasks are related to each other and therefore the timing unit supplies control information to most of the functional blocks in order to ensure proper operation.
The stability of the timing generator is based on a crystal oscillator.

Error detection and correction principles

While playing a disc, disturbances may occur which cause distortion of the audio information. These disturbances manifest themselves in the form of a tick.
They may be the result of scratches, dust or dirt on the disc.
Analogue systems barely (if at all) allow correction of these disturbances.
Digital systems, however, offer the possibility of correction in a relatively simple way.
To enable correction it is necessary to take certain measures prior to recording the codes on the disc.

Single error detection
To create the possibility of detection and/or correction of errors, redundancies must be introduced.
When a code word e.g. consists of 4-bits, there are $2^4 = 16$ possible code word values.
If all these code word values are used for information transmission, no value is left to detect whether the code word received is containing erroneous bits.
The addition of one bit, a so called parity bit, to the 4-bit words, allows to discover the occurence of a single error in the code word.

Suppose we are in an even parity scheme. This implies that, in case the number of 1s in the 4-bit word is even, the parity bit added is 0 (the parity bit will be 1 when the number of 1s is odd). Thus, 5-bit words with an even number of 1s are created. Only words with even parity are valid code words.
If - at the receiver side - a word containing an odd number of 1s is received, there must have slipped in an error, because the total number of 1s (including the parity bit) should be even.

Example
4-bit word A A: 1100
 4321

4-bit word B B: 1000
 4321

The number of 1s in word A is even. Using even parity, the parity bit to be added to this word is 0.

A´: 11000
 43210

So, A´ is a 5-bit word.

The number of ones of word B is odd. If a parity bit is added to word B, it will be 1.

B´: 10001
 43210

A´ and B´ have each an even number of ones and have been derived from the code words A and B.

A 5-bit code has 32 different words, where only 16 are used and 16 are not valid.

0000	**0**
0001	**1**
0010	**1**
0011	**0**
0100	**1**
0101	**0**
0110	**0**
0111	**1**
1000	**1**
1001	**0**
1010	**0**
1011	**1**
1100	**0**
1101	**1**
1110	**1**
1111	**0**
	parity symbol

Multiple error detection
To have the possibility of detecting more than one error in a code word, the addition of one single parity bit will not suffice to allow unambiguous detection, for in case two zeros in code word B have flipped to two ones, B´ will still contain an even number of ones.

What will happen when e.g. two errors are introduced in a given code word A ?

A 000111 ⟶ A´ 010101

In that case code word A´ will differ from A in two positions.

A 000111
A´ 010101

Each code word A'_n with two errors relative to A will differ from A in two positions.
All code words A'_n are wrong code words and are therefore not allowed to belong to the valid code words.

Which criteria should be met by valid code words to ensure unambiguous detection when two errors have slipped into these words?

When two errors occur in A, the received code word A'_n differs from A in two positions.
When all valid code words differ from A in three positions, a code word A'_n will never equal one of the valid code words. If all valid code numbers differ from one another in three positions, it will in no case be possible to obtain a valid code word after the introduction of two errors in a given code word.

What applies to two errors will also apply to three errors, on the condition however that the valid code words differ in four positions.

In general applies:
To be capable of detecting t_1 errors in a code word, all of the admitted code words should differ in t_1+1 positions.

The code words A´ and B´ of the example of single error detection comply with this rule for detecting one error, because differing from one another in two positions, viz. bit 0 and bit 3.

If, during transmission, three errors occur in code words which differ in four positions a decoder should be capable of determining unambiguously that the received code word became corrupted, without mistaking it for another code word.

Example
If during transmission, three errors are introduced in code word A, the received code word A´ should never be identical to B, C or D, because A, B, C and D differ in four positions.

Code word A:	000111	B:	011001
C:	101010	D:	110100

All possible code words A'_n containing 3 errors relative to A, differ from A in three positions.

All these code words A'_n differ both from A and from B, C and D; thus, none of the code words A'_n represents a valid code word.

A: 00 01 11
B: 01 10 01
C: 10 10 10
D: 11 01 00

A: 00 01 11
B: 01 10 01

B: 01 10 01
C: 10 10 10

A: 00 01 11
C: 10 10 10

B: 01 10 01
D: 11 01 00

A: 00 01 11
D: 11 01 00

C: 10 10 10
D: 11 01 00

A, B, C and D differ at least in 4 positions.

A: 00 01 11 $\quad d_m \geqslant 4$
B: 01 10 01
C: 10 10 10
D: 11 01 00

A'_1 : 11 11 11
A'_2 : 11 00 11
A'_3 : 11 01 01
A'_4 : 11 01 10
A'_5 : 10 10 11
A'_6 : 01 10 11
A'_7 : 00 10 01
A'_8 : 00 10 10
A'_9 : 00 11 00
A'_{10}: 00 00 00
A'_{11}: 01 01 00
A'_{12}: 10 01 00
A'_{13}: 10 11 01
A'_{14}: 10 00 01
A'_{15}: 10 11 10
A'_{16}: 10 00 10
A'_{17}: 01 11 01
A'_{18}: 01 00 01
A'_{19}: 01 11 10
A'_{20}: 01 00 10

All possible code words A'_n

If a maximum of three errors is introduced in a code word and the minimum distance between the valid code words is greater than or equal to four, a decoder is capable of determining unambiguously by means of comparison whether or not a received code word is wrong.
To the code words A, B, C and D applies that $d_m = 4$.

$$d_m = t_1 + 1$$

By the minimum distance (d_m) for a number of code words is meant: the minimum number of positions where all valid code words differ from each other.

Error correction
There exists also a possibility to correct errors. This requires, however, that the code words differ from each other in more positions than required for error detection.

To be capable of correcting an error the error-correction circuit should know which code words are correct ones.
The circuit is containing, as it were, a list of correct code words.
Each time a code word enters, it is compared with those contained in the list. If the received word does not figure in the list, the code must be wrong.
After having found that the received code word is wrong, a further comparison is performed to see in how many positions the received code is differing from the correct code words.
The correct code word that differs from the received one in the lowest number of positions is considered to be the word that was transmitted.

The requisite for a proper correction is that a wrong code is unambiguously related, that is, it belongs to one single correct code.
Wrong code A´ of correct code A is not allowed to be equal to wrong code B´ of correct code B. The wrong code A´ and B´ must differ from each other in at least one position.
When A´ and B´ are created after the introduction of 1 error in A and B, A´ and B´ differ in one position from A and B respectively.
To create the possibility of correction, A and B must differ from each other in three position, for A´ will then differ from B in two positions, whereas B´ is differing from B in one position only.

A :	00 00 00
A_1':	00 00 01
A_2':	00 00 10
A_3':	00 01 00
A_4':	00 10 00
A_5':	01 00 00
A_6':	10 00 00
B :	00 11 01
B_1':	00 11 00
B_2':	00 11 10
B_3':	00 10 01
B_4':	00 01 01
B_5':	01 11 01
B_6':	10 11 01

If A´ is differing from B at least in two positions and B´ is differing from B in one position, A´ and B´ will differ from each other in at least one position.

In the Figure $A_1´$, $A_3´$ and $A_4´$ differ from B in two positions.
While: $A_1´$ differs in one position from $B_3´$ and $B_4´$
$A_3´$ differs in one position from $B_1´$ and $B_4´$
and $A_4´$ differs in one position from $B_1´$ and $B_3´$.

In general applies:
To be capable of correcting t_2 errors, the minimum distance should comply with:

$$d_m = 2t_2 + 1$$

To correct one error d_m should at least be 3.

Example
A 6-bit code A (00 00 00) is transmitted; A is one of the valid 6-bit codes with a minimum distance of 3. Due to some disturbance, the following code appears at the receiver side:

A´: 00 10 00

If we compare this code with D and H we find that they differ from each other in three positions;
when compared with C and G in four positions;
when compared with B and F we find a difference in two positions
when compared with E we find a difference in five positions.

However, when compared with A: 00 00 00, the code received differs in one position only. A must therefore be the code that has been transmitted. So the erroneous code must be replaced by code A, which corrects the error.

A:	00 00 00	
B:	00 11 01	
C:	01 01 10	
D:	01 10 11	
E:	10 01 11	
F:	10 10 10	$d_m \geqslant 3$
G:	11 00 01	
H:	11 11 00	

A: 00 00 00
A´: 00 10 00

A: / A´:	00 **00** 00 / 00 **10** 00	d = 1
B: / A´:	00 **11** **01** / 00 **10** **00**	d = 2
C: / A´:	**01** **01** **10** / **00** **10** **00**	d = 4
D: / A´:	**01** 10 **11** / **00** 10 **00**	d = 3
E: / A´:	**10** **01** **11** / **00** **10** **00**	d = 5
F: / A´:	**10** 10 **10** / **00** 10 **00**	d = 2
G: / A´:	**11** **00** **01** / **00** **10** **00**	d = 4
H: / A´:	**11** **11** 00 / **00** **10** 00	d = 3

(d = difference)

If the valid code words only differ in two positions and an error is introduced during transmission, it is possible to detect this error.
The received code word will then differ from the transmitted code word in one position, but there will be several code words that differ from the received word in one position only.

The following code words differ from each other in at least two positions ($d_m \geqslant 2$).

If code word C is the word that is transmitted, but - due to a disturbance - received with one erroneous bit, e.g. C′: **1**101, it is quite easy to detect that C′ is an incorrect code; however, it is not possible to establish unambiguously which one of the bits is wrong, for C′ differs in one bit position from code words C, E, G and H.
If C′ is received, the decoder is not able to determine whether the correct word should be C, E, G or H.
Conclusion, correcting one erroneous bit in code words with a minimum distance $d_m \geqslant 2$ is not possible.

A:	00 00	
B:	00 11	
C:	01 01	
D:	01 10	
E:	10 01	
F:	10 10	$d_m \geqslant 2$
G:	11 00	
H:	11 11	

C: 01 01
C′: 11 01

A:	**00 00**	$d = 3$
C′:	**11 01**	
B:	**00 1**1	$d = 3$
C′:	**11 0**1	
C:	**0**1 01	$d = 1$
C′:	**1**1 01	
D:	**0**1 **10**	$d = 3$
C′:	**1**1 **01**	
E:	1**0** 01	$d = 1$
C′:	1**1** 01	
F:	1**0** **10**	$d = 3$
C′:	1**1** **01**	
G:	11 0**0**	$d = 1$
C′:	11 0**1**	
H:	11 **1**1	$d = 1$
C′	11 **0**1	

Detection and correction of symbols in fr es

In the preceding chapters we learned that it is possible to detect and correct errors in a binary word consisting of a number of bits.

To allow detection the following formula applies:

$$d_m = t_1+1$$

To enable correction the formula to be applied is:

$$d_m = 2t_2+1$$

In the Compact Disc system, however, the samples are converted into 16-bit words. Each 16-bit word is split up into two symbols of 8-bits prior to being passed on to the e)ders.

i he encoder 24 of such symbols are combined in one frame.

Since-in the Compact Disc system frames are formed by symbols, it is quite logical to examine the possibility of detecting and correcting wrong symbols in a frame, instead of looking at the different bits and detecting or correcting wrong bits.
This possibility exists.
In order to detect t_1 wrong symbols in one frame, in all valid frames at least t_1+1 symbols must differ.

So: $d_{ds} = t_1+1$
(d_{ds} = distance to detect wrong symbols)

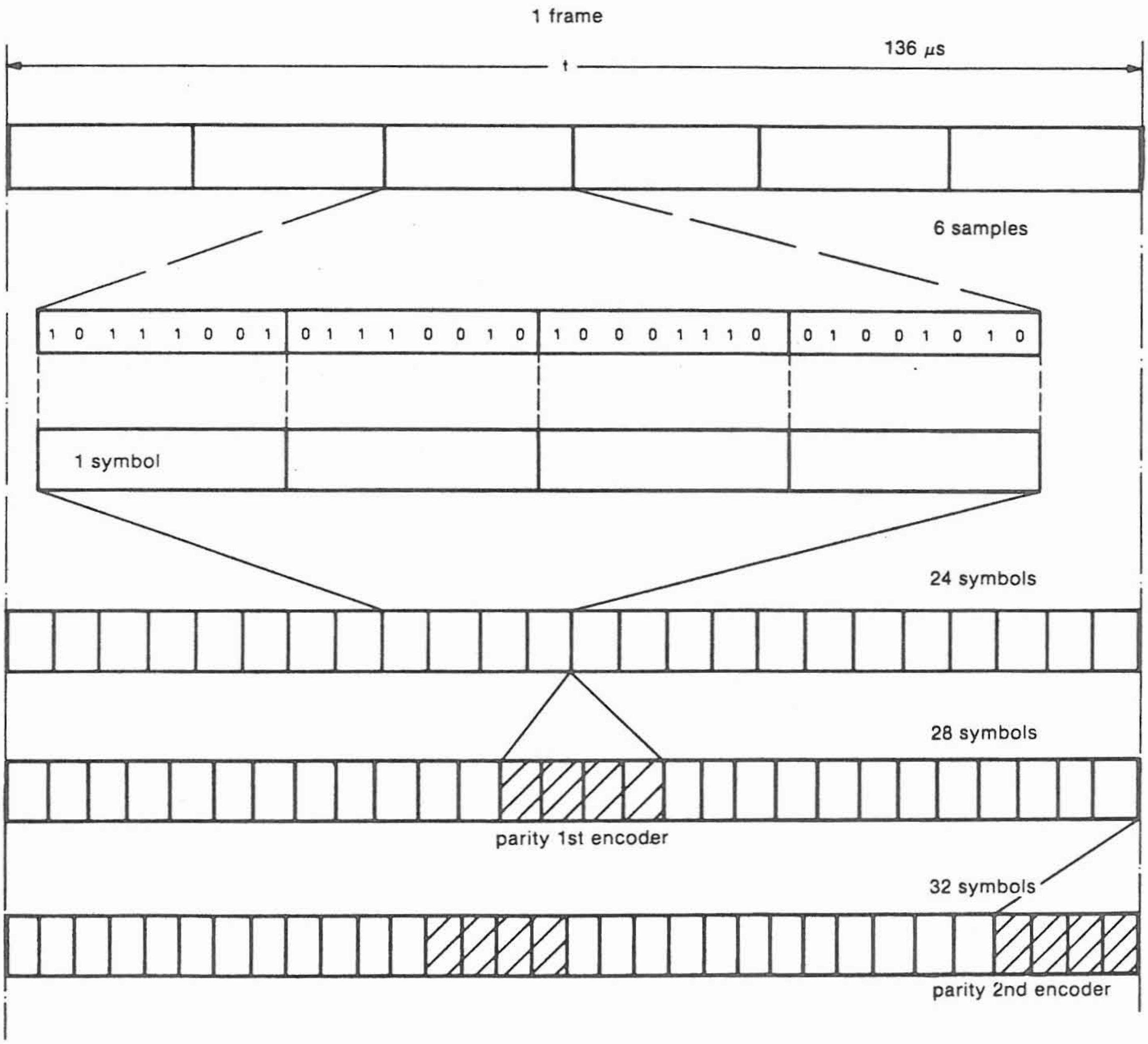
T :orrect t_2 wrong symbols, the frames ɔuld differ in at least $2t_2+1$ positions,

so: $d_{cs} = 2t_1+1$
(d_{cs} = distance to correct wrong symbols)

The same equations that were derived in the previous sections appear to have their validity in this case also.
To comply with these equations it suffices to add parity symbols.
In the Compact Disc system, every batch of 24 audio symbols is complemented with 8 parity symbols in the encoder, thus bringing the number of symbols in 1 frame to 32.
In fact, the error-correction symbols are added to the audio symbols in two distinct encoders; the first adds 4 parity symbols which are related to 24 symbols, wheras the s ɔnd encoder adds 4 parity symbols which

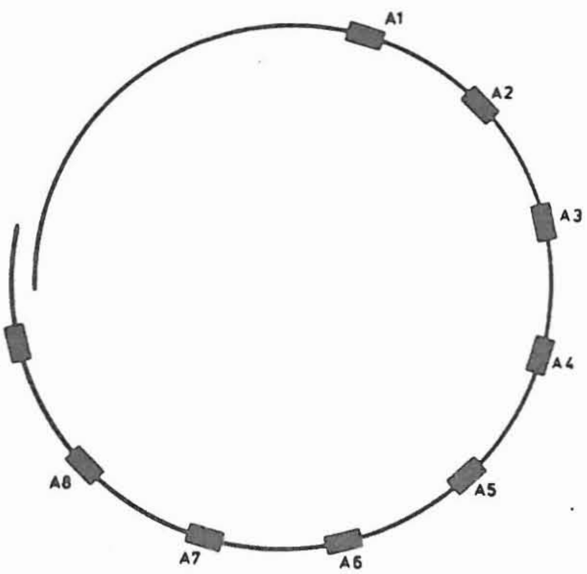
are related to the 28 symbols coming from the first encoder.
The minimum distance for the frames will thus be: $d_{ms} \geqslant 5$ (both for the 24 symbols and the 28 symbols).
(d_{ms} = minimum distance for symbols)

In the Compact Disc player's error-correction IC the parity symbols are placed in a matrix and the information of a frame is multiplied by this matrix.
The result of these multiplications consists of four equations, called syndromes.
These equations make it possible to detect and correct errors.



Scrambling and Interleaving

The majority of errors which may occur during playback of a disc will result from scratches, dust and dirt.
Consequently, each time such a defect occurs a number of adjacent symbols on the disc will be read erroneously.
If all of the affected symbols belonged to the same frame, a great many multi-position errors would occur inside each word and, consequently, correction would be out of the question.
This situation can be avoided by giving up grouped recording on the disc of the symbols belonging to the same code word and by adopting a spread recording by weaving them according to a fixed pattern with symbols from other code words; this process is referred to as interleaving of symbols.



As a result of this, the second symbol of frame A (A_2) will be recorded on the disc after a fixed number of symbols belonging to other frames. If an error burst occurs so that a complete string of symbols is read erroneously, it will nevertheless be possible to correct the errors, because the affected symbols belong to various different frames. In the Compact Disc system it is possible to correct up to 7 frames, that is, 7 blocks of 32 symbols, thanks to the application of this interleaving system.

Interleaving is brought about by means of delay lines having different delay times and allocated to specified symbols.
During encoding the interleaving between the two encoders is taking place. As a result of the interleaving the symbols to which the 1st encoder has added parity symbols will constitute a frame with completely different symbols for which the parity symbols are determined by the second encoder.
The first symbol A_1 coming from the first encoder goes without delay to the second encoder; the second symbol A_2 arrives 4 symbols later, the third A_3 eight symbols later, the fourth A_4 twelve symbols later, and so on.
The fifth symbol A_5 will constitute a frame with Q_1, M_2, I_3 and E_4 for the second encoder.

If it proves impossible to correct an occurring error according to the method described above, there exists a second method to reconstruct a value that approximates the correct value as closely as possible. This method is referred to as interpolation method.

Error detection and error correction of symbols

In the Compact Disc system the samples are converted into 16-bit words. Each 16-bit word is split up into two symbols of 8 bits prior to being passed on to the encoders.
To every batch of 24 symbols four parity symbols are added in the first encoder, thus giving a total of 28 symbols.
In a second encoder another group of 4 parity symbols is added to the previous total of 28 symbols, thus bringing the number of symbols to 32.
A string of 32 symbols is called a frame.
Since - in the Compact Disc system - the words are formed by symbols, it is quite logic to examine the possibility of detecting and correcting wrong symbols, instead of looking at the different bits and detecting/correcting wrong bits.
This possibility exists. In order to detect t_1 wrong symbols, the symbols representing the various code words positions, should differ from each other in a minimum of $t_1 + 1$
so: $d_{ds} \geqslant t_1 + 1$.
(d_{ds} = minimum distance to acted erroneoms symbols).

To correct t_2 wrong symbols, the symbols should differ in $2t_2 + 1$ positions, so:
$d_{cs} = 2t_2 + 1$.

Once we know which symbol is wrong, it is even possible to correct simultaneously a greater number of wrong symbols.
A prerequisite is that a decoder-independent device indicates which symbols are wrong.
The maximum number of symbols that can be corrected is then:
$t_{cs} = d_m - 1$.

(d_{ms} = min. dist. of symbols)

Example

The following code words consisting of three symbols composed of two bits differ from one another in three positions.

A 00 00 00	$d_{ms} = 3$
B 01 01 01	$d_{ms} = t_1 + 1 \quad t_1 = 2$
C 10 10 10	$d_{ms} = 2t_2 + 1 \quad t_2 = 1$

If it is known which symbols are wrong, we get: $t_2 = d_{ms} - 1 = 2$

It must therefore be possible to detect $t_1 = 2$ wrong symbols, to correct directly $t_2 = 1$ error and, if we know which symbols are wrong, to correct $t_2 = 2$ symbols.

A: 00 00 00
B: 01 01 01
C: 10 10 10
D: 11 11 11 $\quad d_{ms} = 3$

Assume that code A is transmitted and that - at the receiver side - one symbol is wrong, ȷ. X: 11 00 00.
Comparison with the words B, C and D reveals a difference in 3, 3 and 2 positions respectively.
Only comparison with A shows a difference of 1.
A must therefore be the correct code word.

A: 00 00 00
X: 11 00 00

A: 00 00 00
X: 11 00 00 $d_s = 1$

B: 01 01 01
X: 11 00 00 $d_s = 3$

C: 10 10 10
X: 11 00 00 $d_s = 3$

D: 11 11 11
X: 11 00 00 $d_s = 2$

It is possible to detect and correct one erroneous symbol is $d_{ms} \geqslant 3$.

If the code word received contains two wrong symbols, e.g. Z: 11 11 00, in no case a valid code word can be formed;
from this follows that a detector whose operation is restricted to error detection only is capable of discovering two errors.
If, however, it is known that code Z_0 contains two wrong symbols and, furthermore, which symbols are wrong, it is possible to correct the error.
By replacing the wrong symbols by all possible symbols,

A: 00 00 00
Z: 11 11 00

A: 00 00 00
Z: 11 11 00 $d_s = 2$

B: 01 01 01
Z: 11 11 00 $d_s = 3$

C: 10 10 10
Z: 11 11 00 $d_s = 3$

D: 11 11 11
Z: 11 11 00 $d_s = 1$

It is possible to detect two erroneous symbols if $d_{ms} \geqslant 3$
Correction **not** possible

We must necessarily find the correct word among the words thus formed.

Z_0: 11 11 00
 3 2 1

Symbols 2 and 3 are wrong; replace 2 and 3 by all possible combinations and compare these code words with the valid ones.

Comparison of Z_0 through Z_{15} with the valid code words A, B, C and D yields the unambiguous result that Z_{15} is the same as A. So, the correct code word must be A.

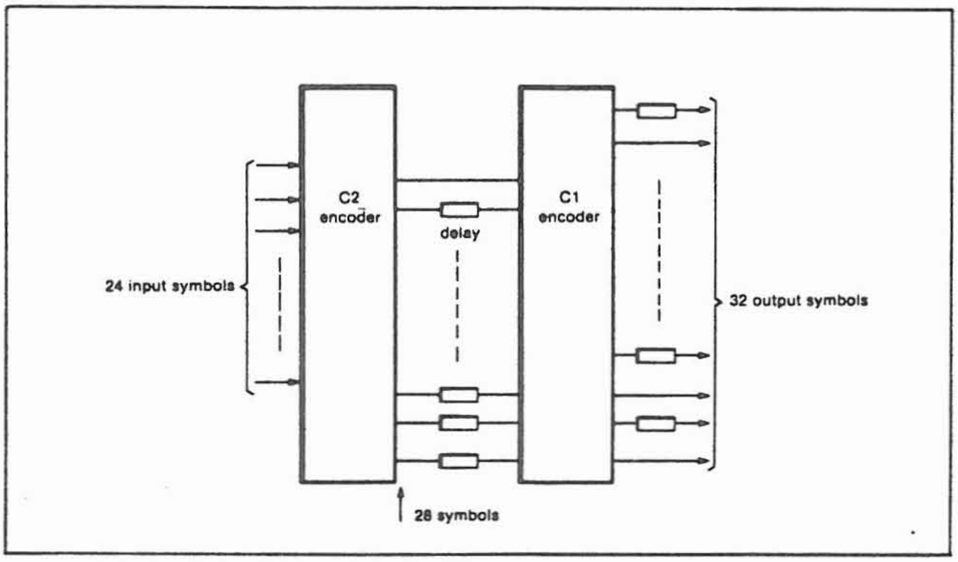
This allows the conclusion that - once we know which symbols in a code word are wrong - it is possible to correct $t_2 = d_{ms} - 1$ errors.

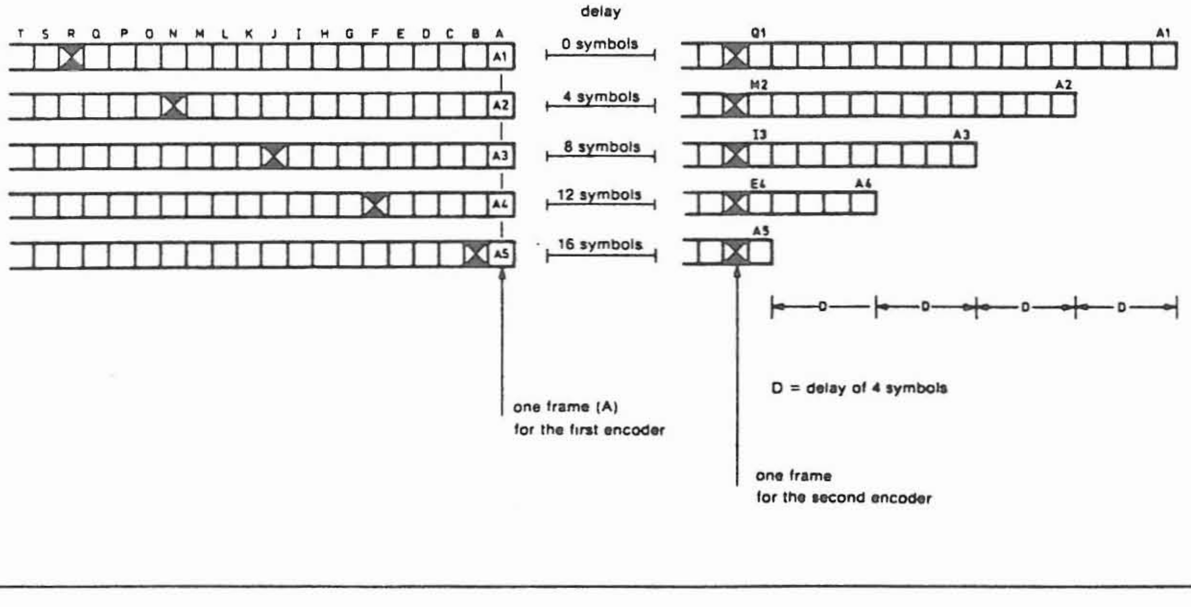
The Compact Disc system has been designed to correct in the C_2 decoder errors which have been detected by the C_1 decoder, because the C_1 decoder adds a flag, that is, an error signal, to each defect symbol.

Z_0 :	11 11 00
Z_1 :	11 10 00
Z_2 :	11 11 00
Z_3 :	11 00 00
Z_4 :	10 11 00
Z_5 :	10 10 00
Z_6 :	10 01 00
Z_7 :	10 00 00
Z_8 :	01 11 00
Z_9 :	01 10 00
Z_{10}:	01 01 00
Z_{11}:	01 00 00
Z_{12}:	00 11 00
Z_{13}:	00 10 00
Z_{14}:	00 01 00
Z_{15}:	00 00 00

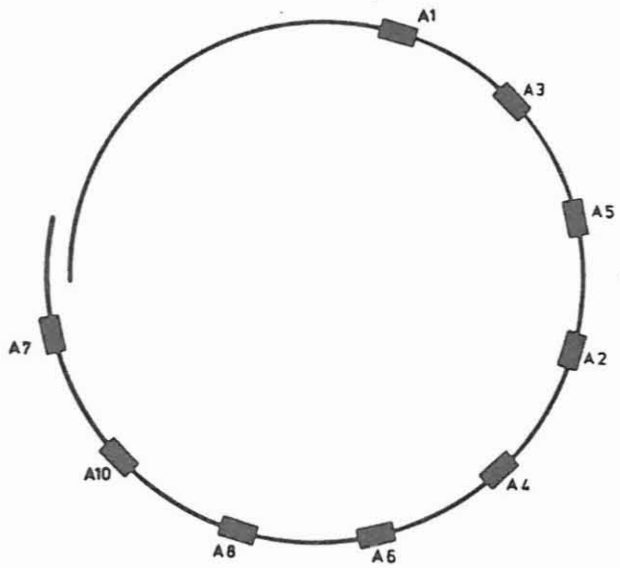
A: 00 00 00
B: 01 01 01
C: 10 10 10
D: 11 11 11 $d_{ms} = 3$

Analogue signals are continuous signals which will usually not change abruptly. The amplitude of the signal during the first sample will not differ greatly from that during the second sample. The amplitude during the third sample will not differ much from the second, and so on.
If the value of the second sample cannot be corrected and the values of the first and the third samples are known, a good approximation must allow to compute the value of the second sample.





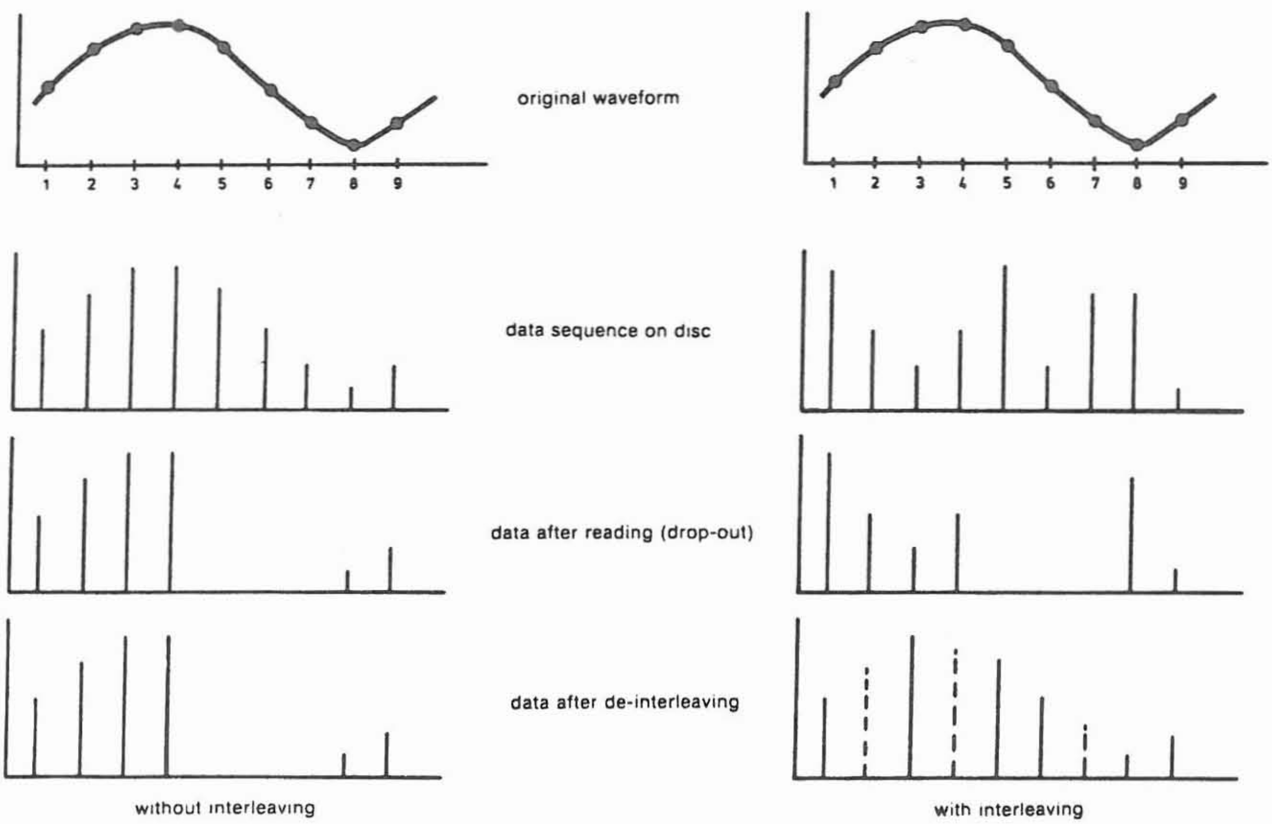
To enable the application of this method, the code word symbols are interchanged crosswise so that reading of symbol A_1 (of code word A) is followed by e.g. A_3 (instead of A_2) and then A_5; this is called: Scrambling of symbols.
If a disturbance causes erroneous reception of A_3 and A_5 and correction is impossible, we still have the symbols A_1, A_2, A_4 and A_6 which allow to compute interpolated value of A_3 and A_5.



The principle of interleaving is illustrated in the Figure.

In the left-hand part, the sequence of signal processing tasks is shown without interleaving.
The audio signal is first sampled at time points 1, 2, 3, etc. than digitized and recorded on the disc.

If there is a drop-out during the reading of the disc, there will be some symbols missing in the received data. In the example used, three symbols are missing.
As the length of the drop-out region is greater than the error correction capability of the decoder, there is no chance of reconstructing the missing values, i.e. the audio output has to be muted in order to avoid clicks.

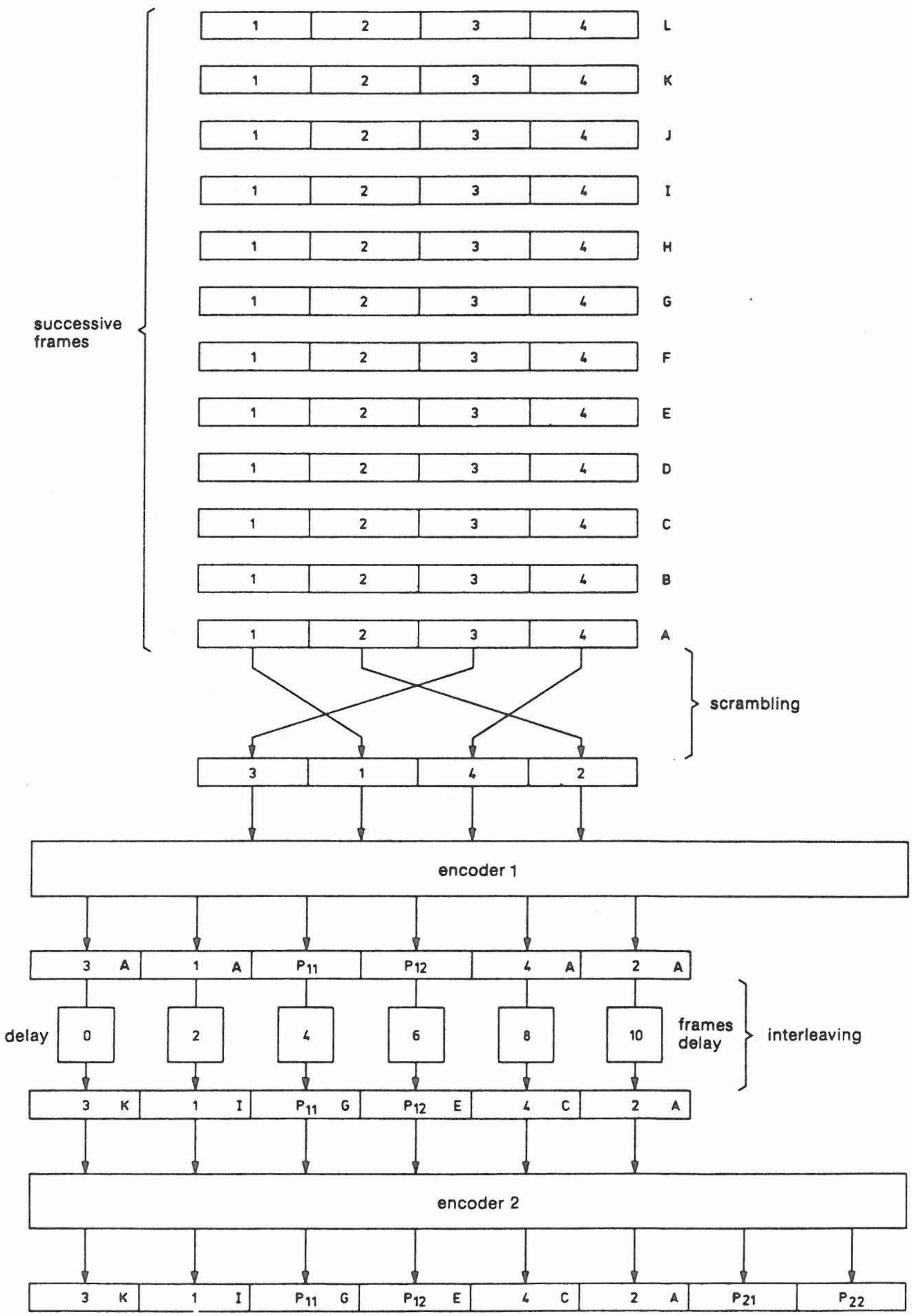


In the right-hand part of the Figure the same sequence of signal processing activities is s wn, this time with interleaving of data. The original audio signal is again sampled, this time with the samples being re-arranged prior to disc recording. This interleaving results in the recording of data in a sequence which does not represent an increasing time scale. During the reading of the disc, the same drop-out is assumed, again resulting in three missing symbols. Then, de-interleaving is performed to restore the original sequence of the data symbols.
Now it can be seen, that the drop-out region has been 'spread out' over a larger period of time and that there are only single error inputs to the decoder. They can be corrected easily. The dashed lines represent the recovered data.

Example of scrambling and interleaving for four symbols

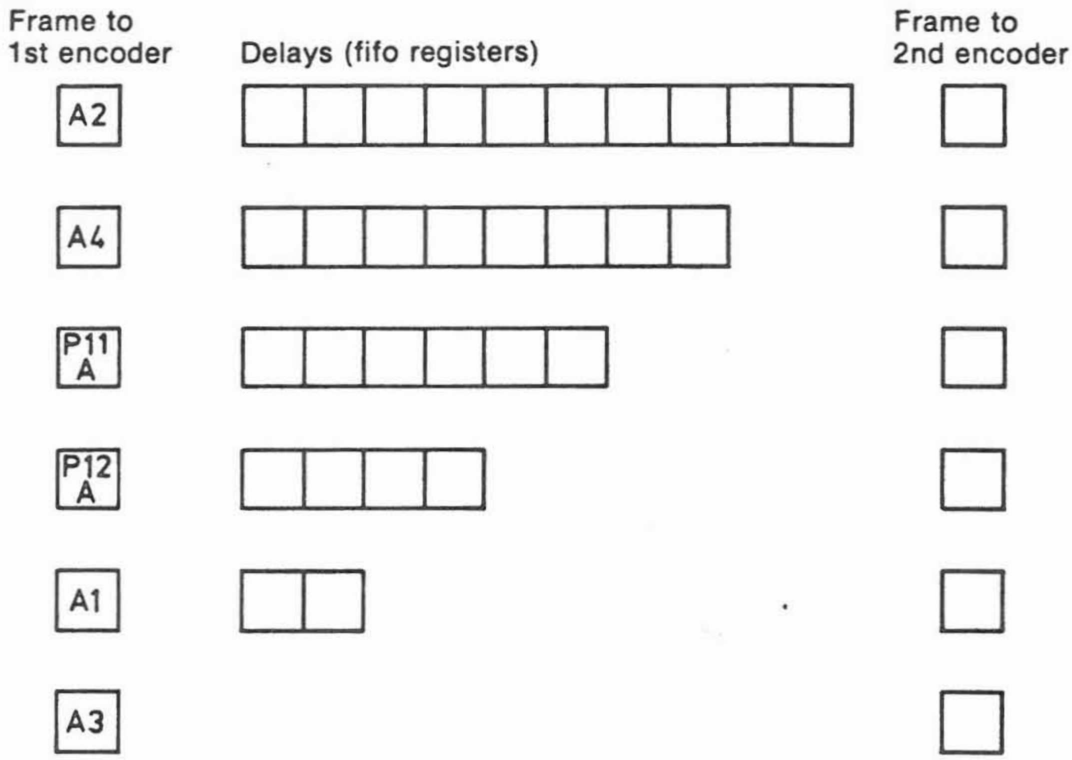
The picture shows 12 successive frames of four symbols as they are offered to the encoding circuitry, all symbols of one frame in parallel. To the frames first scrambling is applied; this means that the order of the symbols in the frames is changed. The first encoder adds (in this case) two parity symbols to the frame, P11 and P12 (in the compact disc four parity symbols are added by the first encoder).
The following action taken is interleaving. Interleaving is obtained by giving each symbol a differend delay before it is offered to the second encoder. The delay in this case for the first symbol (symbol 3) is zero, symbol 3 is not delayed in this case. Symbol 1 however is delayed for two frames, symbol 4 for eight frames and symbol 2 for ten frames. Also the parity symbols get a different delay P11 four frames and P12 six frames.
The frame offered to the second encoder will consist out of:
Symbol 2 of frame A
Symbol 4 of frame C
Symbol 1 of frame I
Symbol 3 of frame K.
The parity symbols in this frame offered belong originally to:
P11 to frame G
P12 to frame E.
The second encoder now calulates again two parity symbols, P21 and P22 these symbols are added to the frame too.
The delays can be compared with first-in first-out shift registers of different length.
The following figures give an example of the interleaving delays for a frame of four symbols and two parity symbols.
In the figures the frames are drawn in vertical direction as they come from the first encoder and are offered to the interleaving delays.
The horizontal bloks represent the delays of resp. 10,8, 6,4 2 and 0 frames.
A clockpulse shifts the symbols into the registers each clockpulse the symbols of the next frame. After eleven clockpulses the first complete frame is offered to the second encoder, consisting of the following symbols: A2, C4, P11E, P12G, I1, K3.
Compare the figure.

Example of scrambling and interleaving for 4 symbols

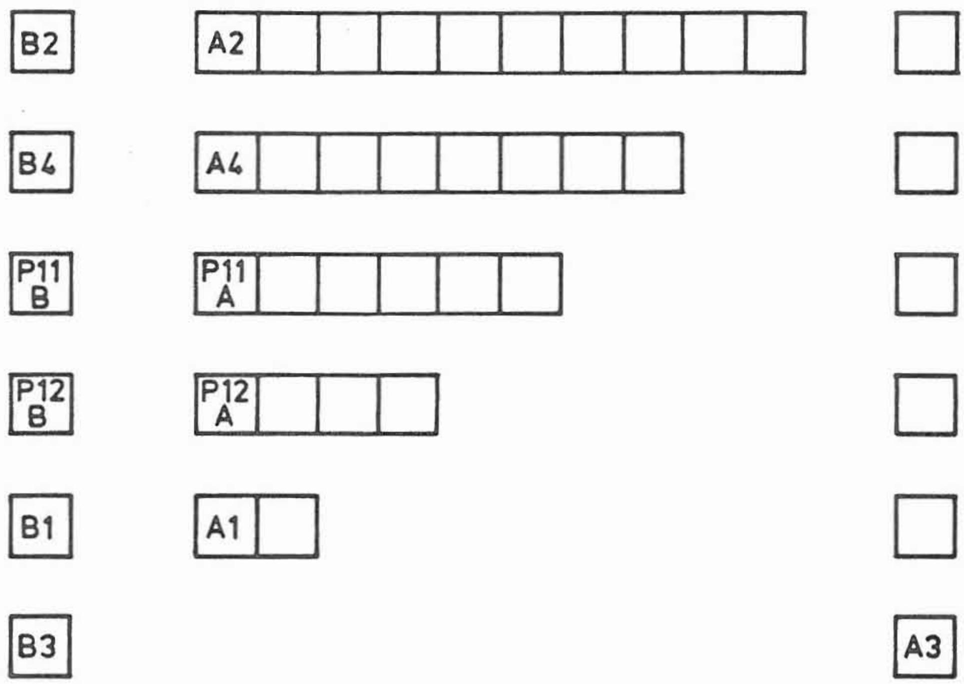


Note: P_{11}, P_{12} parity symbols encoder 1
P_{21}, P_{22} parity symbols encoder 2

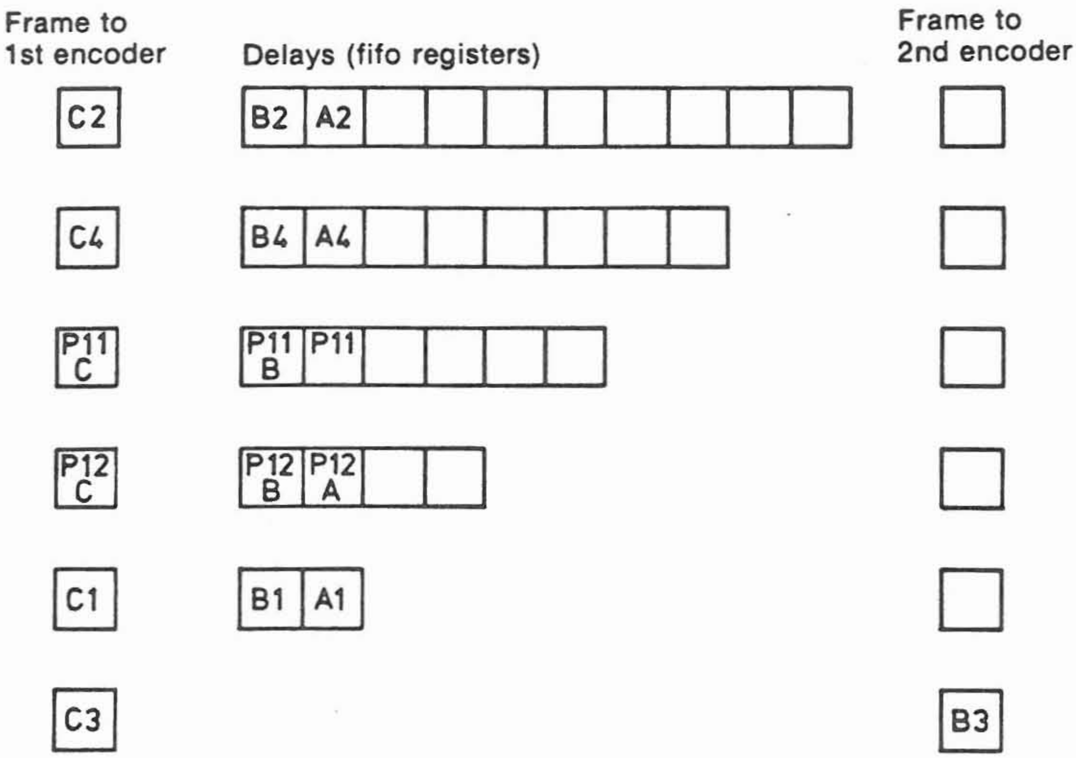
Initial state



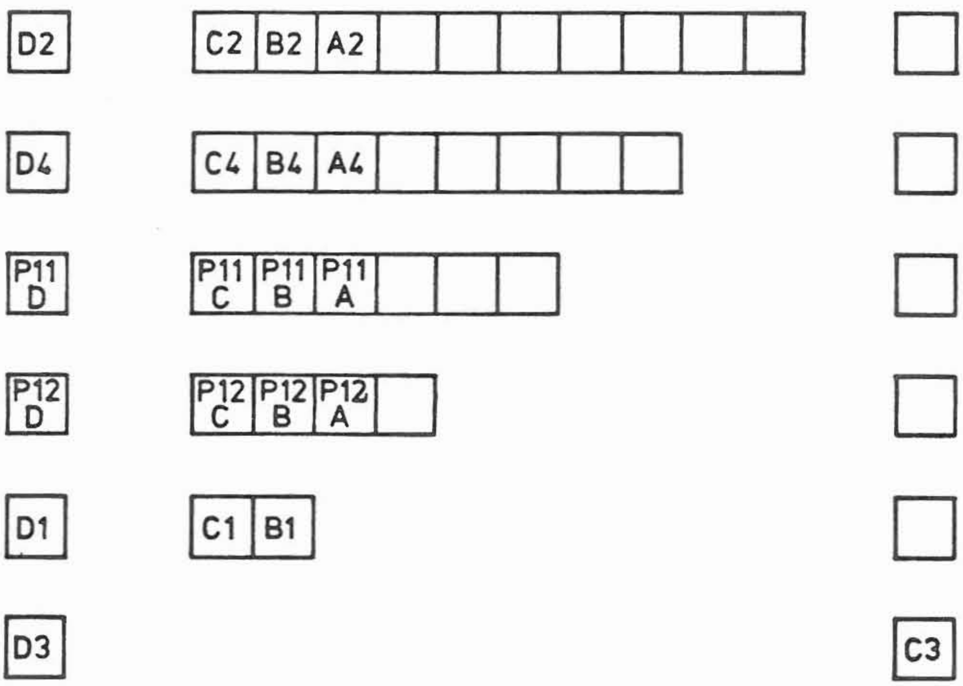
State after one clock pulse



State after two clock pulses



State after three clock pulses



State after four clock pulses

Frame to 1st encoder	Delays (fifo registers)										Frame to 2nd encoder
E2	D2	C2	B2	A2							
E4	D4	C4	B4	A4							
P11 E	P11 D	P11 C	P11 B	P11 A							
P12 E	P12 D	P12 C	P12 B	P12 A							
E1	D1	C1									B1
E3											D3

State after five clock pulses

Frame to 1st encoder	Delays (fifo registers)										Frame to 2nd encoder
F2	E2	D2	C2	B2	A2						
F4	E4	D4	C4	B4	A4						
P11 F	P11 E	P11 D	P11 C	P11 B	P11 A						
P12 F	P12 E	P12 D	P12 C	P12 B							P12 A
F1	E1	D1									C1
F3											E3

State after six clock pulses

Frame to 1st encoder	Delays (fifo registers)										Frame to 2nd encoder
G2	F2	E2	D2	C2	B2	A2					
G4	F4	E4	D4	C4	B4	A4					
P11 G	P11 F	P11 E	P11 D	P11 C	P11 B	P11 A					
P12 G	P12 F	P12 E	P12 D	P12 C							P12 B
G1	F1	E1									D1
G3											F3

State after seven clock pulses

Frame to 1st encoder	Delays (fifo registers)										Frame to 2nd encoder
H2	G2	F2	E2	D2	C2	B2	A2				
H4	G4	F4	E4	D4	C4	B4	A4				
P11 H	P11 G	P11 F	P11 E	P11 D	P11 C	P11 B					P11 A
P12 H	P12 G	P12 F	P12 E	P12 D							P12 C
H1	G1	F1									E1
H3											G3

State after eight clock pulses

Frame to 1st encoder	Delays (fifo registers)	Frame to 2nd encoder
I2	H2, G2, F2, E2, D2, C2, B2, A2, ,	
I4	H4, G4, F4, E4, D4, C4, B4, A4	
P11 I	P11 H, P11 G, P11 F, P11 E, P11 D, P11 C	P11 B
P12 I	P12 H, P12 G, P12 F, P12 E	P12 D
I1	H1, G1	F1
I3		H3

State after nine clock pulses

Frame to 1st encoder	Delays (fifo registers)	Frame to 2nd encoder
J2	I2, H2, G2, F2, E2, D2, C2, B2, A2,	
J4	I4, H4, G4, F4, E4, D4, C4, B4	A4
P11 J	P11 I, P11 H, P11 G, P11 F, P11 E, P11 D	P11 C
P12 J	P12 I, P12 H, P12 G, P12 F	P12 E
J1	I1, H1	G1
J3		I3

State after ten clock pulses

Frame to 1st encoder	Delays (fifo registers)										Frame to 2nd encoder
K2	J2	I2	H2	G2	F2	E2	D2	C2	B2	A2	
K4	J4	I4	H4	G4	F4	E4	D4	C4			B4
P11 K	P11 J	P11 I	P11 H	P11 G	P11 F	P11 E					P11 D
P12 K	P12 J	P12 I	P12 H	P12 G							P12 F
K1	J1	I1									H1
K3											J3

State after eleven clock pulses

Frame to 1st encoder	Delays (fifo registers)										Frame to 2nd encoder
L2	K2	J2	I2	H2	G2	F2	E2	D2	C2	B2	A2
L4	K4	J4	I4	H4	G4	F4	E4	D4			C4
P11 L	P11 K	P11 J	P11 I	P11 H	P11 G	P11 F					P11 E
P12 L	P12 K	P12 J	P12 I	P12 H							P12 G
L1	K1	J1									I1
L3											K3

State after twelve clock pulses

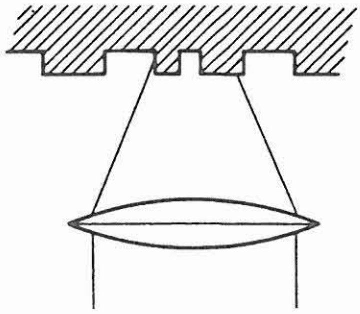
Frame to 1st encoder	Delays (fifo registers)	Frame to 2nd encoder
M2	L2, K2, J2, I2, H2, G2, F2, E2, D2, C2	B2
M4	L4, K4, J4, I4, H4, G4, F4, E4	D4
P11 M	P11 L, P11 K, P11 J, P11 I, P11 H, P11 G	P11 F
P12 M	P12 L, P12 K, P12 J, P12 I	P12 H
M1	L1, K1	J1
M3		L3

State after thirteen clock pulses

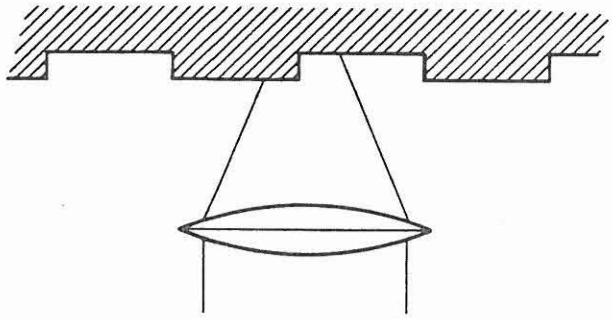
Frame to 1st encoder	Delays (fifo registers)	Frame to 2nd encoder
N2	M2, L2, K2, J2, I2, H2, G2, F2, E2, D2	C2
N4	M4, L4, K4, J4, I4, H4, G4, F4	E4
P11 N	P11 M, P11 L, P11 K, P11 J, P11 I, P11 H	P11 G
P12 N	P12 M, P12 L, P12 K, P12 J	P12 I
N1	M1, L1	K1
N3		M3

Modulation

A modulation system for an optical audio disc must fulfil certain requirements:
- self-clocking ability
- read-out at high information density
- low spectral power at low frequencies
- small error propagation



As the bit clock has to be re-generated from the data after the reading of the disc, self-clocking ability is mandatory. This is normally achieved by making the maximum distance between high-to-low and low-to-high transitions in the data as small as possible.



The light spot which is used for the reading of the disc has finite dimensions. This could lead to inter-symbol interference because of the high information density used on the disc. To avoid this , the minimum distance between transitions in the data should be as large as possible.

Low spectral power at low frequencies is desirable because DC and low frequencies could lead to interference with the servo systems.

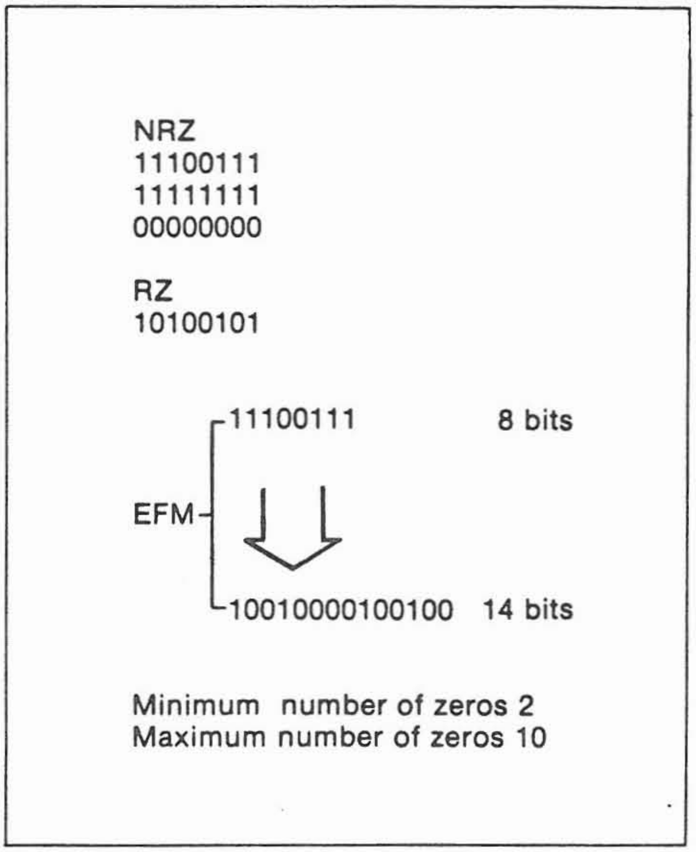
The propagation of errors should be kept as small as possible in order to restrict their influence and to limit the amount of hardware needed for the correction of errors.

As data which are presented in a non-return-to-zero format (NRZ), are not self-clocking and may contain a large DC component, they are not suitable for optical disc recording.

The afore-mentioned requirements are met by an encoding procedure termed Eight-to-Fourteen Modulation (EFM). As the name implies, each group of 8 data bits from the error correction encoding unit (audio and parity bits) or from the control and display encoding unit is encoded into 14 channel bits.

In order to ensure self-clocking ability as well as to permit the reading of high-density information, there are always at least two 0's (zeros) between successive 1s (ones) and no more than ten zero's in a run.

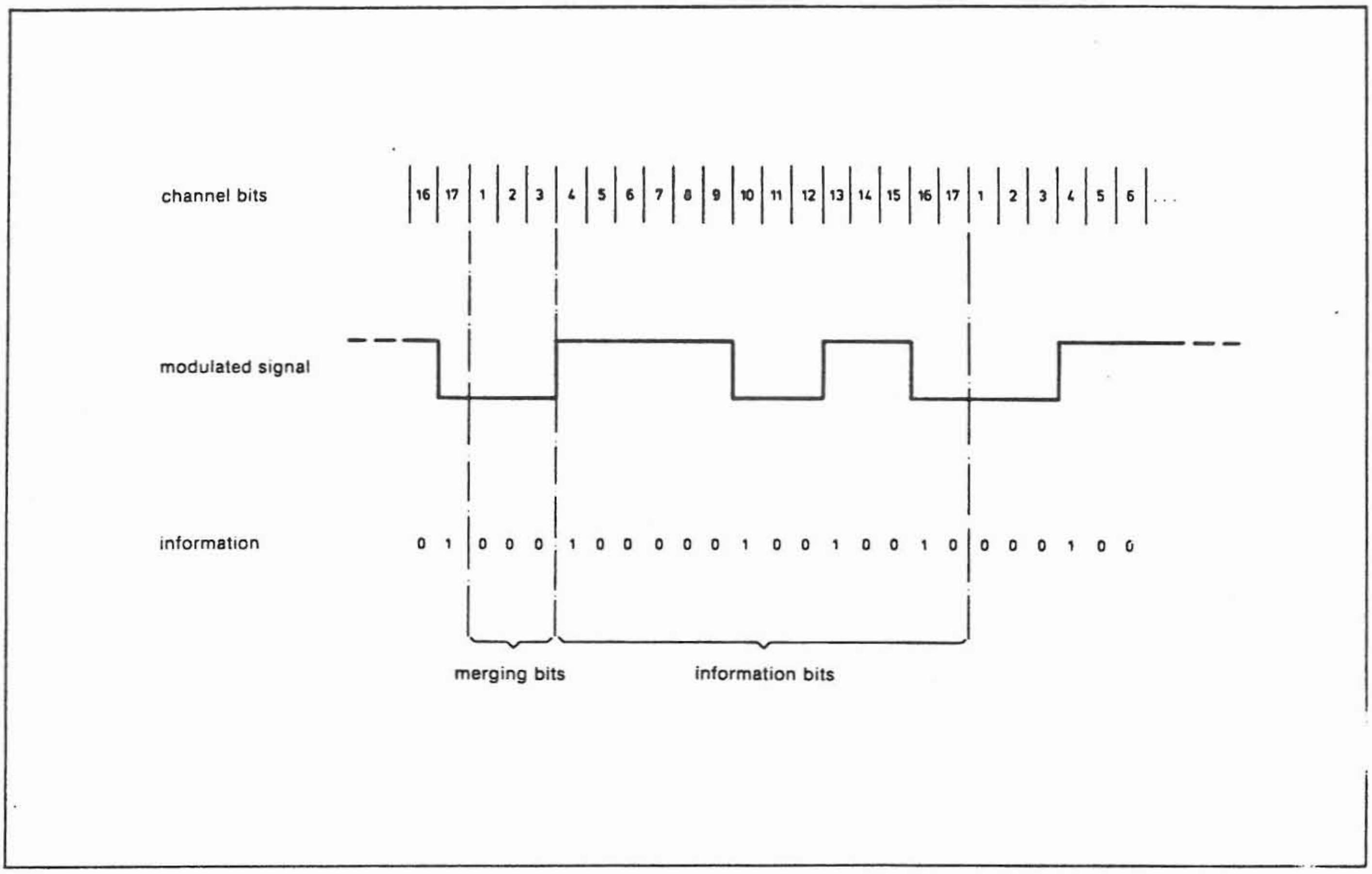
There exist 277 different 14-bit sequences which satisfy the constraint of at least two zero's and maximum ten.
Deleting the 21 patterns with the longest run-length leaves 256 sequences which perfectly match the $2^8 = 256$ possibilities of the 8-bit wide input data.
Thus, there is a one-to-one correspondence between 8-bit NRZ data and 14-bit EFM data. A code conversion can now easily be achieved by means of a look-up table which can be stored in a ROM.
The 14-bit blocks generated in this way cannot, however, be concatenated without violating the constraint of the two 0s at the block boundaries.
In order to solve this problem, three merging bits are inserted between successive blocks. These merging bits do not contain any audio or display information and are therefore skipped by the decoder. They are only used to comply with the contraints of the two to ten 0s and to control the DC content of the resulting modulation stream.
DC control is achieved by inserting or omitting an extra transition in the merging bits. The decision is based upon the knowledge of one or more future symbols (look-ahead method). An example of the modulation signal is given in the Figure. The transmitted information is contained in the position of the transitions.



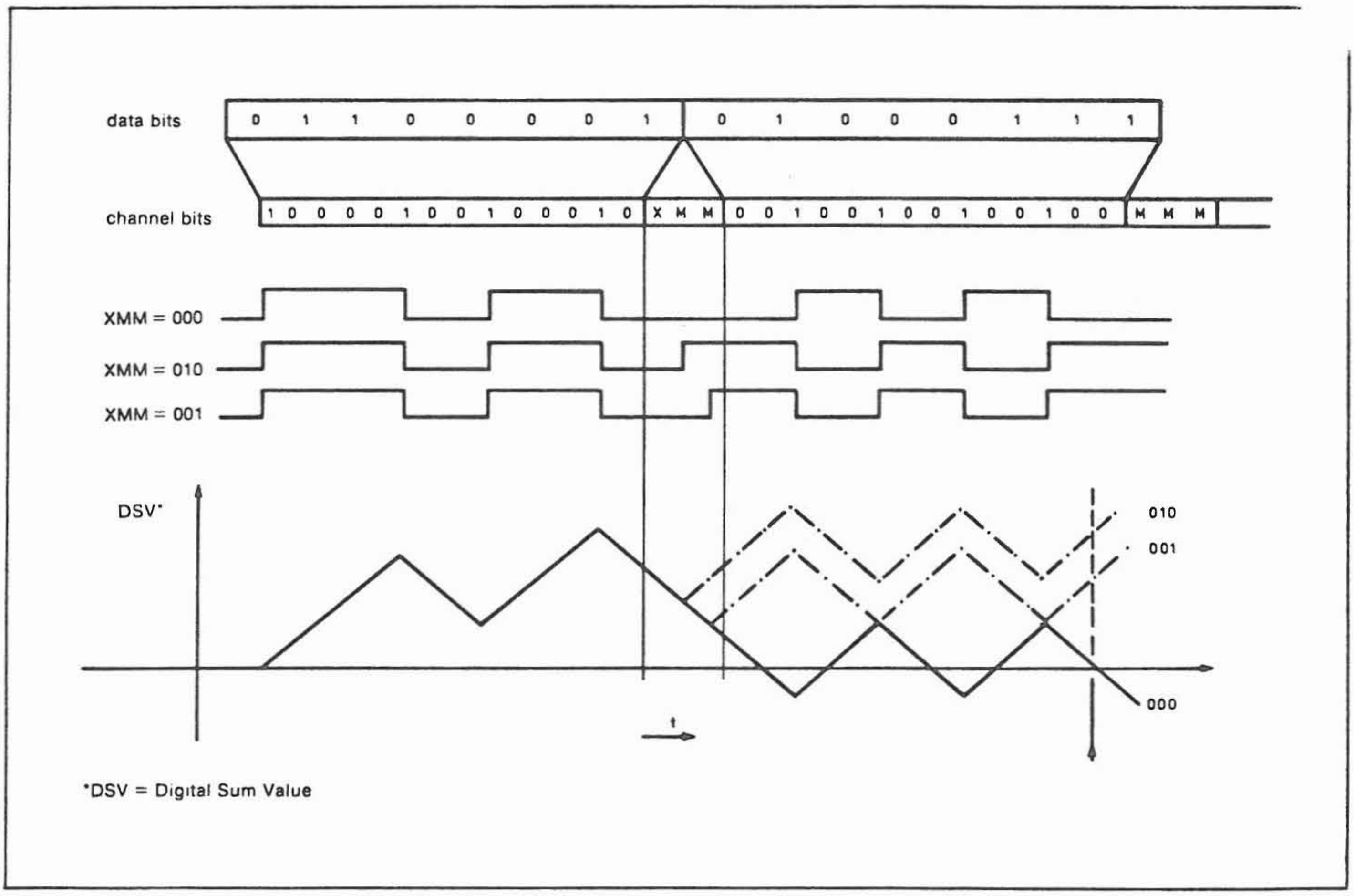
As EFM is based upon a block structure of 8-bit wide input data, it is very well suited for the adopted CIRC error correction system which is also based on blocks of 8 consecutive data bits. The propagation of errors is limited to the 8 data bits forming a symbol.

Due to the requirement of self-clocking ability some kind of synchronization is necessary. This is achieved by dividing the data stream into frames and adding a unique pattern to each frame which does not occur in the data.

Each frame contains:
- a special synchronization pattern
- 12 data words of 16-bits each
- 4 error-correction parity words of 16 bits each
- a control and display symbol of 8-bits



channel bits
16 17 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 1 2 3 4 5 6 ...
modulated signal
information
0 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 1 0 0 0 0 1 0 0
merging bits
information bits

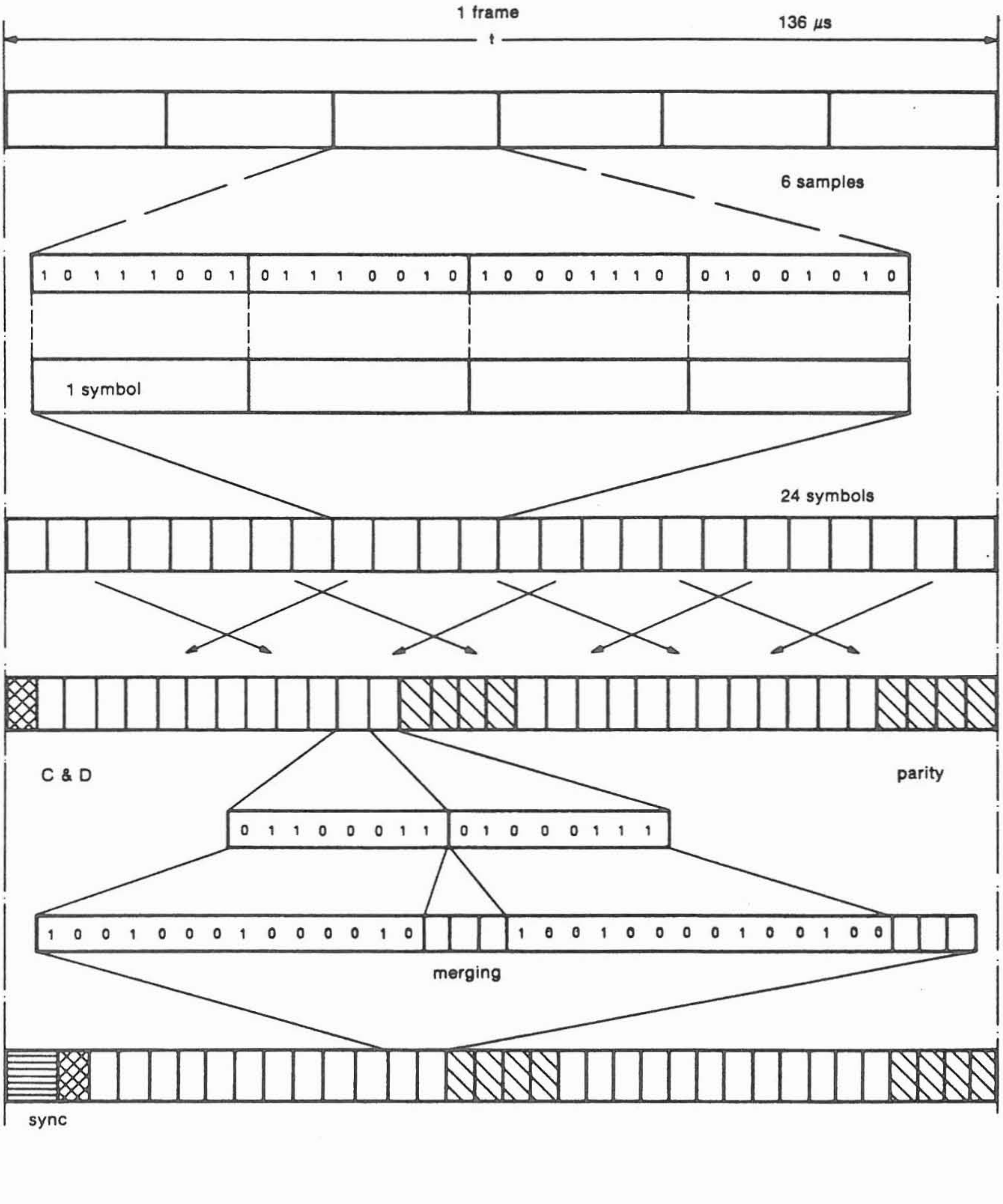


data bits
0 1 1 0 0 0 0 1 0 1 0 0 0 1 1 1
channel bits
1 0 0 0 0 1 0 0 1 0 0 0 1 0 X M M 0 0 1 0 0 1 0 0 1 0 0 1 0 0 M M M
XMM = 000
XMM = 010
XMM = 001
DSV*
010
001
000
t
*DSV = Digital Sum Value

The data and parity words are each split-up into two 8-bit blocks before EFM encoding es places.

The total number of channel bits per frame is:

- sync pattern		24 channel bits
- control and display	1 x	14 channel bits
- data	12 x 2 x	14 channel bits
- error correction	4 x 2 x	14 channel bits
- merging and DC control	34 x	3 channel bits
Total		588 channel bits



Example of one frame of 4 samples

In the figure is shown how the binary information of one frame of four samples results in a track of pits on a disc.
For simplicity reasons scrambling and interleaving is left out.
The first string of binary information represents four samples of audio information of 16 bits.
For reasons of error-correction and easyer handling of the information the 16 bits samples are splitted up into two symbols of 8 bits each.
The next string of information therefore is formed by 8 symbols of 8 bits.
The codes of the symbols are up to now normal binary codes, which vary from 00000000 to 11111111.
Putting this information on the disc in the way described before, two problems show up:
- if a symbol contains all zero's, no pits will be pressed on the disc.
- the lengthes of the pits and the distances between the pits will be of the same order as the diameter of the laserspot.

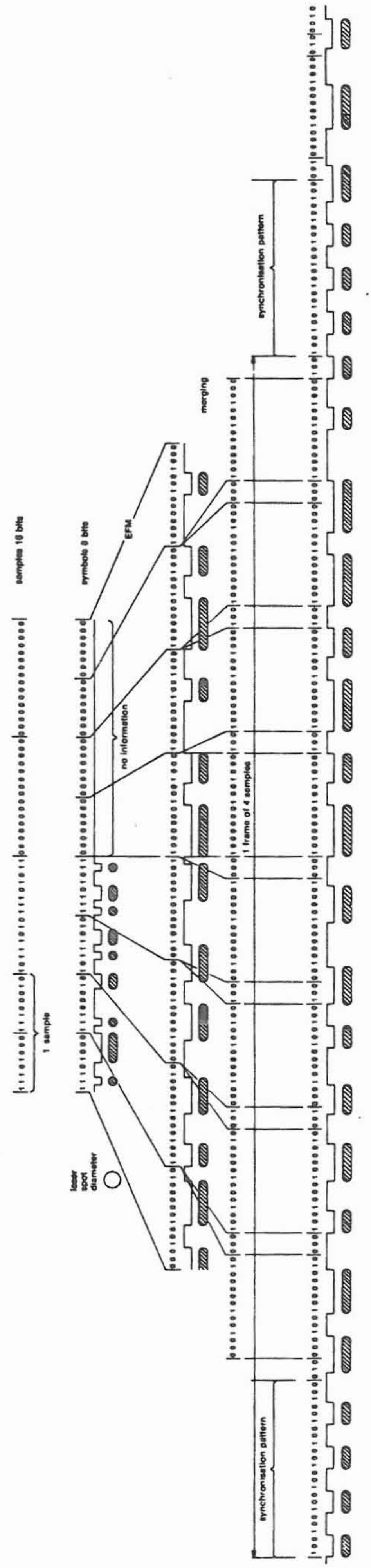
If no pits are available tracking is not possible. If pit-lengthes and distances are equal to the laserspotdiameter, the reflected laserlight will be modulated by more than one pit at a time, which gives inter pit interference.

To overcome these problems, in the compact disc use is made of the EFM modulation technique.
By EFM each 8 bits symbol is converted into a symbol of 14 bits. The 14 bits symbols now meet the requirement that between two one's at least two and maximum ten zero's are placed.

The result of this is:
- codes with all zero's are converted to codes containing zero's and one's.
- the lengthes and distances of the pits are at minimum three bits.

The lengthes and distances of the pits is defined for the symbols now, but at the boundaries of the symbols it can happen that these lengthes still are too short.
For this reason and for the reason of the DC content three merging bits are introduced between two 14 bits symbols.

In order to make it possible for the player to know where astring of new information begins, every, frame starts with a unique patern of one's and zero's the sync-pattern.



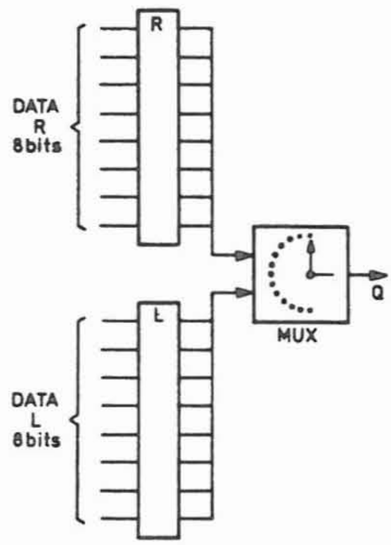
Example of one frame of 4 samples

Time conversion

w additional information should be added to audio information in order no to do violence to the continuity of the final audio signal.

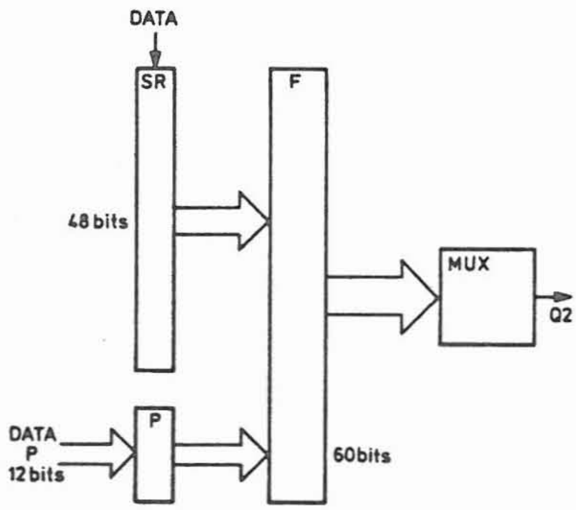
Example

Assume that every second two 8-bit registers are loaded with data.
This implies that every second the content of the registers is renewed.
A multiplexer scans the parallel outputs of the registers at a speed of 16 bits per second. The multiplexer will then have read the last bit at the moment that new data is clocked in so that that the multiplexer can proceed further from the first bit onwards.



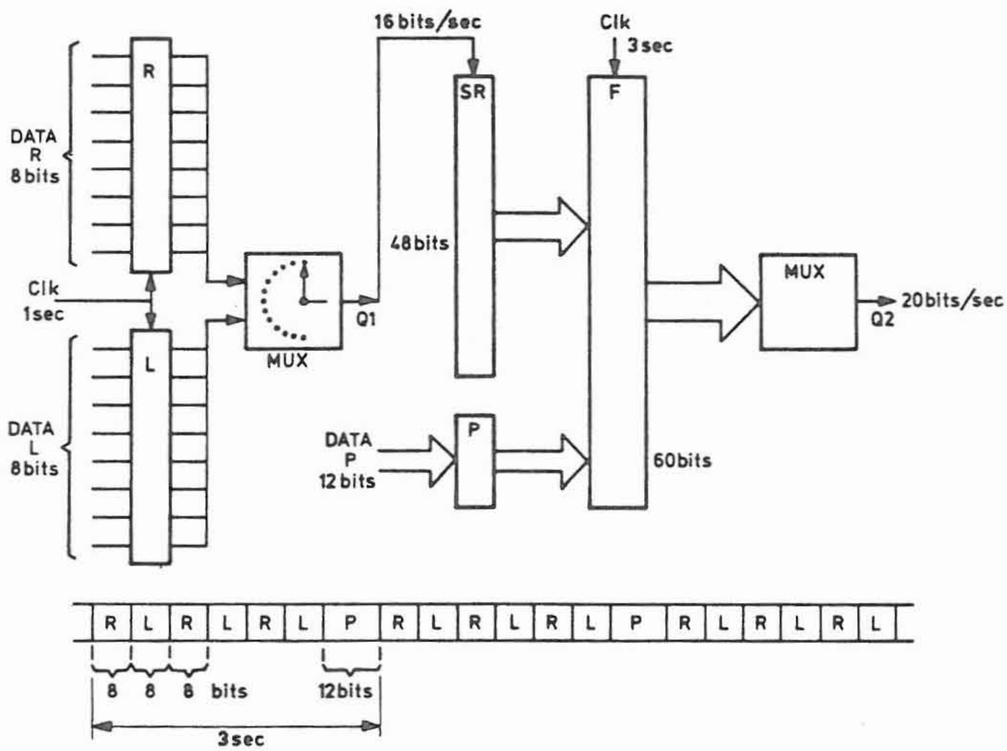
The DATA coming from the multiplexer may be moved into a shift register capable of aining e.g. 48 bits.
hen the multiplexer has scanned three es either of the registers the shift register will be filled; at the moment that the shift register is filled, the DATA (48 bits) is parallel loaded into another register. This register is provided with new data every 3 seconds.
It is also possible to clock DATA from the shift register and from another source simultaneously into the register, e.g. DATA present in another register that can hold e.g. 12 bits.

The register that receives the DATA should, of course, have a storage capacity of
48 + 12 = 60 bits.



If the multiplexer reads the 60 bits in 3 seconds, that is, at a speed of 20 bits per second, it will just have finished this job when the DATA in the register is renewed, so that the multiplexer can restart from the beginning.
The Figure below shows the total diagram of this system.
In this way a continuous data flow is obtained that contains 3x the information that was present in either of the two 8-bit registers completed with the DATA from the 12-bit register.
This may be schematically represented as follows:

P R L R L R L P



In the Compact Disc system every 22,7 μs (44,1 kHz) a sample is taken of both the right and the left channel.
Via an Analogue-to-Digital Converter these samples are expressed in 16-bit binary codes.
Every 22,7 microseconds these samples are clocked in registers (R and L).
By means of a multiplexer these registers are read out and their content is moved into a shift register.

This shift register has the capacity to hold 12 samples, that is, 192 bits.
the moment that the 12 samples are present in the register, the parity information is also available in 8 symbols of 8 bits, that is, in total 64 bits.
The 192 bits of audio information and the 64 bits of parity information are clocked simultaneously into another register that has a capacity of 246 bits.
This register is loaded every 12 x 22,7 = 272,4 μs.
A multiplexer takes 272,4 μs to read out this register, thus maintaining, a continuous flow of data that contains both audio and parity information.

In a similar way other types of information, like the sync, control and display codes, may be added to the bit flow.

In the player the reverse process takes place, ıt is, new strings of binary audio formation arrive every 22,7 μs at the ˈgital-to-Analogue Converter.

Appendix B

The DA converter

Questions

1. *What principle is used by the DAC to convert a digital word into an analogue signal?*

2. *Of which basic elements is the DAC composed?*

3. *How is the digital signal finally converted into an analogue signal?*

Digital-to-Analogue Conversion at Playback of the Compact Disc

Introduction

The last section in the sequence of processings to which the signal in the Compact Disc system is subjected, is the re-transformation of the digital code into an analogue signal that has the same shape as the original acoustic vibration that was captured by the microphone.

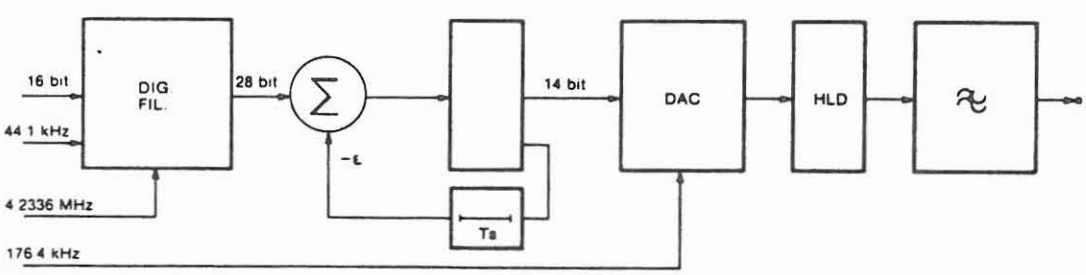
The digital signal takes - after decoding and error correction - the form of a series of 16-bit words.
Each word represents in binary form the numerical quantity of the sound magnitude at that moment and is, therefore, a sample of the audio signal. Each second 44100 of these words appear.

The Digital-To-Analogue Converter (DAC) in the Compact Disc player generates with each word an electric current of associated strength and keeps it constant until the next word arrives. The electric current will thus describe a "stepped" curve that is an approximation to the analogue signal shape. Speaking in term of frequency we may say that the step represent high frequencies which exceed by far the band of the analogue audio signal (20 Hz-20 kHz). These high frequencies have to be suppressed by a low-pass filter (in the Compact Disc system their level has to be brought to at least 50 dB below the maximum audio signal level).

To allow a mitigation of the requirements the low-pass filter should normally satisfy, the approach selected in the Philips Compact Disc player is to perform a preliminary filtering operation during the digital phase. For this, a fourfold oversampling is performed, that is, a digital filter operating at four times the sampling frequency (4 x 44,1 kHz = 176,4 kHz) provides signal values at this raised frequency, thus refining the stepped curve and facilitating the filtering-out of the high frequencies. As a result, it will suffice to use a relatively simple third-order low-pass filter conversion section.

To convert the 16-bit words into an analogue signal the Philips Compact Disc player uses a 14-bit Digital-To-Analogue Converter available in IC form; this DAC is designed to operate at the high sampling frequency of 176,4 kHz. The fourfold oversampling on the one hand and a feedback in phase opposition of the rounding errors on the other hand bring about that the rounding-off to 14 bits does not lead to a degradation of the Signal-To-Noise Ratio (SNR). This SNR is maintained at the level of about 96 dB that belongs to 16-bit quantization.

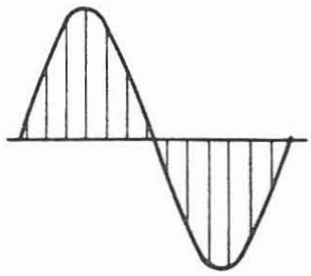
The set of operations belonging to the Digital-To-Analogue Conversion are illustrated in the block diagram.
In connection with stereo reproduction through two channels the complete chain is given in twofold.
Oversampling takes place in the digit filter to which the input signal is fed. The next step consists in rounding off to 14 bits; the rounding error is fed back in phase opposition in the noise shaper (NS).
The digital filter and the noise shaper are accommodated in a single IC realised in NMOS technique.
Then follow the DAC and a hold circuit finally, the analogue signal passes through a low-pass filter.

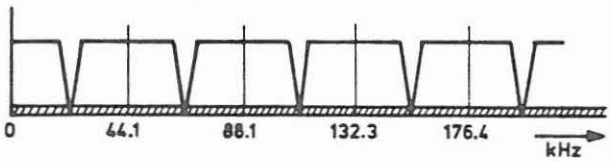


4.2336 MHz = 96 x 44.1 kHz

Suppression of the frequencies above the audio band

Direct digital-to-analogue conversion of the presented signal results in a series of analogue signal samples with a repetition frequency of 44,1 kHz. The frequency spectrum of such a series is shown in the fig. This spectrum is, in priciple, infinitely large; above the base band (0-20 kHz) entire multiples of the sampling frequency with their left and right side bands are shown. Between these bands transistion ranges occur, the first one - for instance - between 20 kHz and 24,1 kHz.

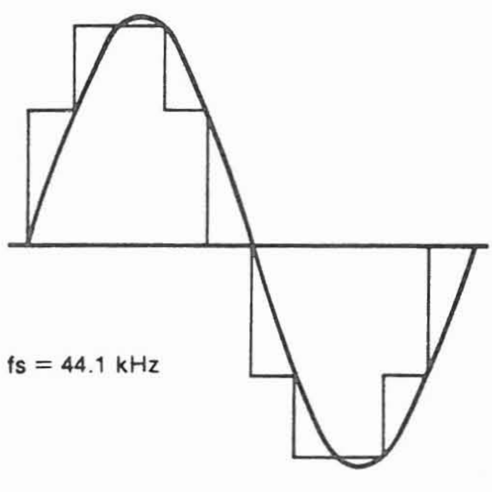


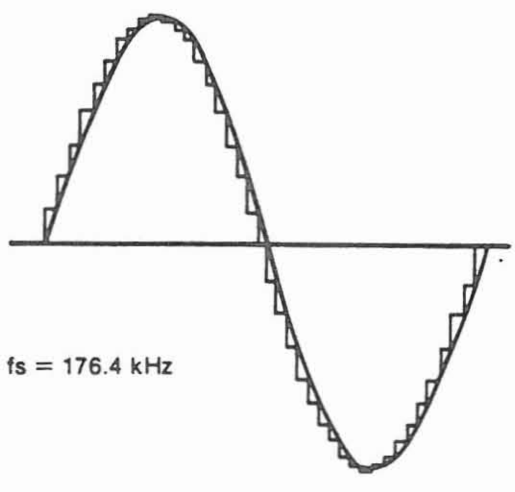


This entire spectrum shall not be passed on to the playback amplifier and the loudspeaker. Even though the frequencies above 20 kHz are inaudible, they might give rise to overload of the playback amplifier and to intermodulation products with the base band frequencies or, for instance, with the high frequency bias current of a tape recorder.
This explains the formulation of the requirement to attenuate all of the frequencies above the base band by at least 50 dB.

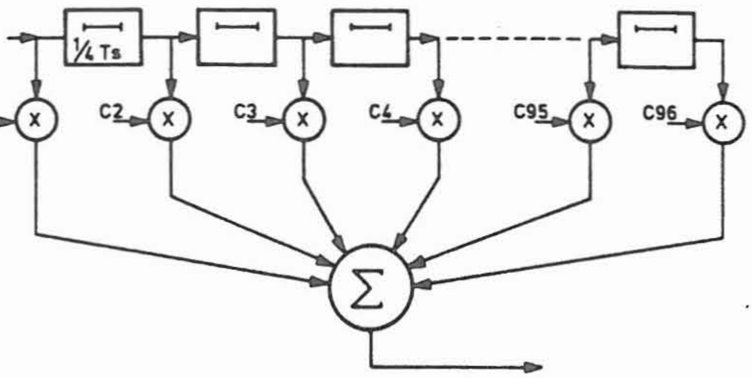
The analogue filter following the DAC must - in order to ensure an attenuation of this importance - inevitably contain a great many elements and require a very accurate adjustment. Furthermore, the pass-band requires a linear phase characteristic in order not to affect the waveshape of pulse-like sound phenomena.
To meet the formulated requirements, a different approach has been adopted in the Philips Compact player, namely:

- a fourfold oversampling of the signal in the digital phase.
- a hold function after the Digital-To-Analogue-Conversion section, and
- a third-order Bessel filter in the analogue signal path.

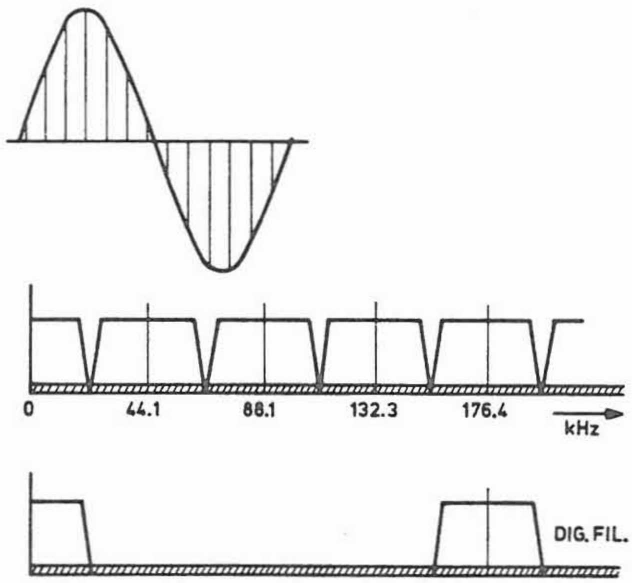




A transversal digital filter comprising 96 elements takes care of the oversampling. The delay in each element is $(176{,}4 \times 10^3)^{-1}$ s, that is, a fourth part of the sampling period. Four times per sampling period the filter takes up new data, but three out of four times its content is nil. Consequently, on the total of 96 elements only 24 filled. The content of each element is multiplied by a coefficient (the coefficients have an accuracy of 12 bits). The filter provides data with the frequency of 176,4 kHz; each numeral is the sum product of 24 mulitplications. In this way the filter interpolates three new sample values between any of two received samples. The numerals supplied have a length of 28 bits; from this follows that no rounding-off occurs in the filter.

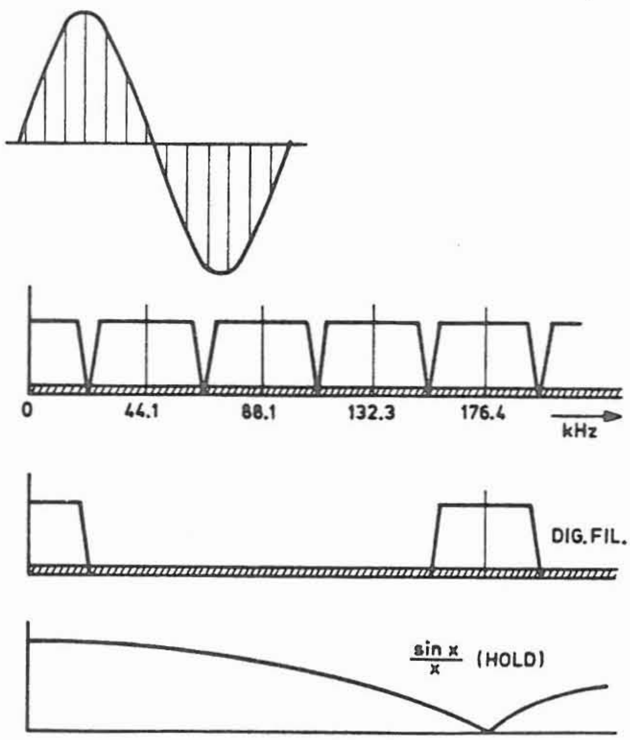


The frequency spectrum of the oversampled signal is illustrated in the Figure. We see that the band around 1x, 2x and 3x 44,1 kHz are no longer present in this spectrum.



The DAC generates a current whose quantity corresponds to the last bit word presented. This current is kept constant in a hold circuit until the next bit word arrives, thus creating the so called stepped curve.
The signal samples thus passed from (in theory) infinite short pulses to pulses with the duration of the sampling period. This has also consequences for the frequency spectrum; the spectrum is multiplied by a curve of the shape (sin x) /x which has a first zero crossing at 176,4 kHz. This results in a considerable attenuation of the frequency band around this frequency.
This attenuation is phase-linear.

It is, however, insufficient. A third-order low-pass Bessel filter, having its —3 dB point at 30 kHz, serves to complete the attenuating action. Bessel filters provide essentially straigth phase characteristics in the pass-band.
This filter is quite simple and requires no precision elements.



The hold function and the Bessel filter introduce a slight damping in the upper part of the pass-band. The digital filter has been designed to correct this damping. The digital filter also corrects the small attenuation resulting from the (sin x) /x effect.
With a view to these corrections the coefficients of the digital filter have been scaled down by approx. 2 dB in order to prevent the occurrence of overload in the accumulator of the digital filter.
This implies that the over-all signal-to-noise ratio is approx. 95 dB.

Example of digital filter with coefficients

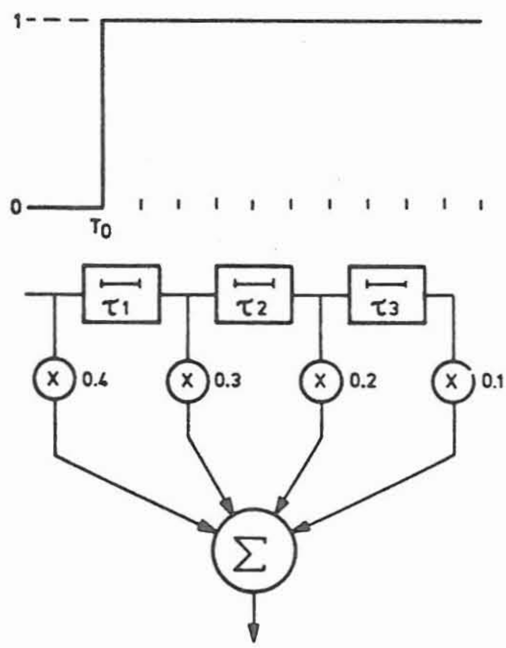
The digital filter drawn has four coefficients: 0.4, 0.3, 0.2, 0.1

If a step function is applied to the input of the filter, the output during T_1 will be:

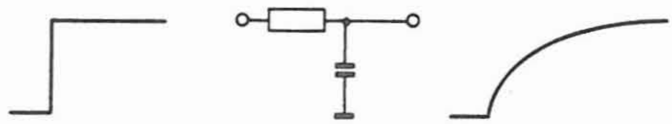
0.4 x 1 = 0.4
0.3 x 0 = 0
0.2 x 0 = 0
0.1 x 0 = 0

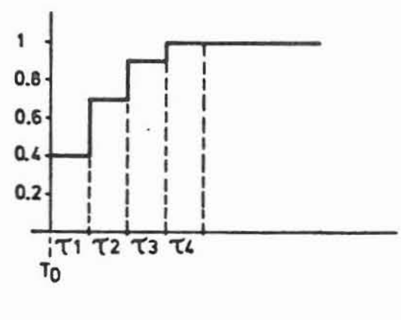
Output 0.4

during T_2	during T_3	during T_4
0.4 x 1 = 0.4	0.4 x 1 = 0.4	0.4 x 1 = 0.4
0.3 x 1 = 0.3	0.3 x 1 = 0.3	0.3 x 1 = 0.3
0.2 x 0 = 0	0.2 x 1 = 0.2	0.2 x 1 = 0.2
0.1 x 0 = 0	0.1 x 1 = 0.1	0.1 x 1 = 0.1
Output 0.7	0.9	1



The step function comprising all frequencies has become a shape rounded off at the edge. If one compares this shape with the output of a R-C filter where the same step is applied to, than we see a simmular shape. The result of the two filters can be considered the same.





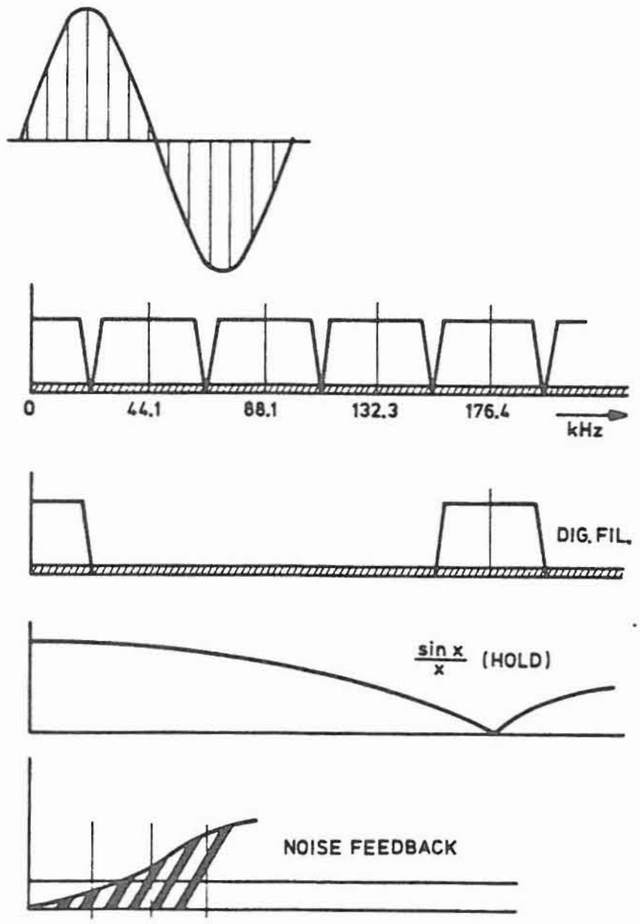
Suppression of the quantization noise

The presented 16-bit quantized signal will - at conversion - be transformed into an analogue signal accompanied by a given noise. This noise results from the errors in the sampled frequency band flowing from the quantization in discrete steps. The effective quantity of the noise voltage in the sampled frequency band is $q/\sqrt{12}$, where q represents the interval size of the quantization step.
We see that at doubling of the quantization step, that is, coding with one less bit, the noise voltage becomes twice as great, or - in other terms - the noise level increases by 6 dB.

When a 16-bit coding scheme, the signal-to-noise ratio is approximately 96 dB. It would be a pity to see this excellent SNR adversely affected by a rounding-off to 14 bits required by the adoption of the 14-bit D/A converter. Without any measures, this rounding-off will lead to a fall in SNR to 84 dB. Fortunately, it is possible to do something about this.

Oversampling does not cange the absolute noise level, since during oversampling no rounding-off takes place and, consequently, the quantization step remains as large as before.
The sampled frequency band has, however, become four times greater. The noise power is uniformly distributed throughout this frequency band (quantization noise is white noise). The band between 0 and 20 kHz is relevant only; this band represents one fourth of the sampled band and the noise power in the band from 0 to 20 kHz also represents one fourth of the over-all noise power. This implies an improvement of the SNR by 6 dB; the SNR thus becomes approximately 90 dB, that is, equal to that of a 15-bit system.

When rounding off from 28 bits to 14 bits it is wise to take advantage of the correlation which exists between successive rounding errors. If the analogue signal is a dc voltage, the successive samples will contain the same rounding error. The audio signal will not contain dc current, but slowly varying signals which- viewed from microscale show resemblance with dc. By adding the error caused during rounding from 28 to 14 bits, in phase opposition to the next sample, it is possible to reduce the average quantization error for slowly varying signals (low frequencies). This is found back in the form of the frequency spectrum at low frequencies the noise level becomes lower, whereas at

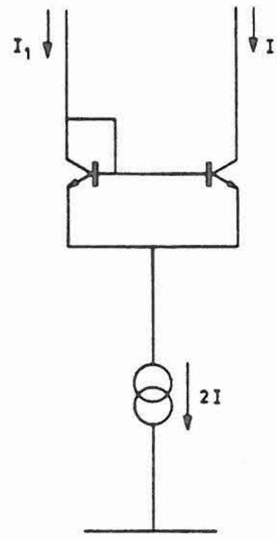


high frequencies it becomes higher. In the audio band we get thus a gain in the SNR of 7 dB. This raises the SNR to approximately 97 dB, which equals at least the value belonging to the initial 16-bit quantization.

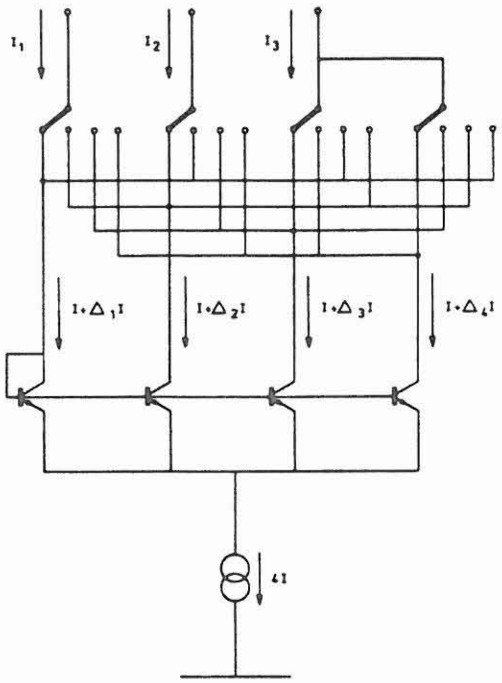
The DA converter

The Philips DA converter converts a digital number at its input into an analogue current proportional to the value of that digital number

The Figure opposite shows the basic circuit for the current sources.
The current 2 I is divided in I_1 and I_2; these currents are almost equal to each other so that only minor deviations $\Delta_1 I$ and $\Delta_2 I$ will occur.
The sum of the currents I_1 and I_2 always equals 2I; consequently, $\Delta_1 I + \Delta_2 I = 0$.

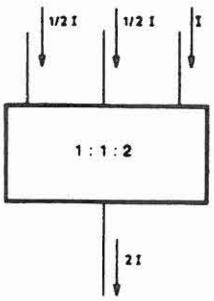


The next Figure represents a circuit in which four transistors are connected similarly.
The collectors of these transistors are, however, connected via a matrix to four switches which, in turn, are connected to three outputs. The switches are actuated by a clock pulse. Each pulse causes the switches to make one step.
After four steps the switches have returned to the position shown in the Figure.



A consequence of this configuration is that the mean currents I_1, I_2 and I_3 are equal to I, I and 2 I respectively, for $\Delta_1 I + \Delta_2 I + \Delta_3 I + \Delta_4 I = 0$.
The proportions of the currents I_1, I_2, I_3 are thus defined very accurately as follows:

$$I_1 : I_2 : I_3 = I : I : 2 I = 1 : 1 : 2.$$



$$I_1 = \frac{4 I + \Delta_1 I + \Delta_2 I + \Delta_3 I + \Delta_4 I}{4} = I$$

$$I_2 = \frac{4 I + \Delta_1 I + \Delta_2 I + \Delta_3 I + \Delta_4 I}{4} = I$$

$$I_3 = 2x\left(\frac{4 I + \Delta_1 I + \Delta_2 I + \Delta_3 I + \Delta_4 I}{4}\right) = 2 I$$

The mean values of the currents are obtained by means of smoothing. The current proportions remain of course valid when the reference current is not 4 I but 2 I, as occurs in the following Figure.
In this Figure the squares with the output current proportions of 1 : 1 : 2 are cascaded. Each following output, starting from I, has a current that is one half of the previous one. In this way it is possible to allot a binary weight to the output currents.

In the next following Figure the outputs of the current sources are connected to switches which interconnect these outputs to the output of the DAC or to the supply voltage. The switches are operated by the binary code that must be converted into an analogue current. In the Figure shown, the binary code 00101110100100 (2980) is presented.

The corresponding current is:

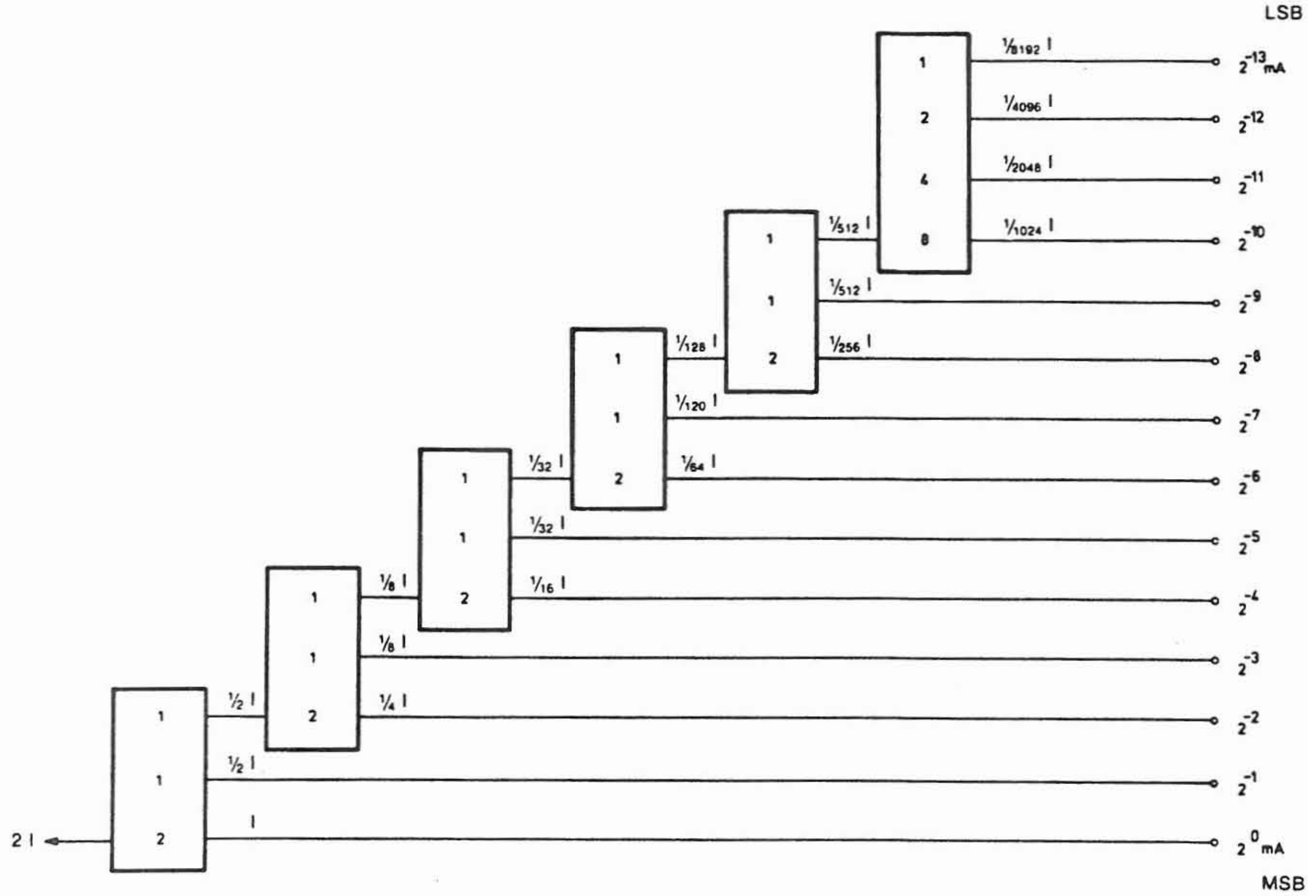
$$\left(\frac{1}{2048} + \frac{1}{256} + \frac{1}{64} + \frac{1}{32} + \frac{1}{16} + \frac{1}{4}\right) I = \frac{744}{2048} I$$

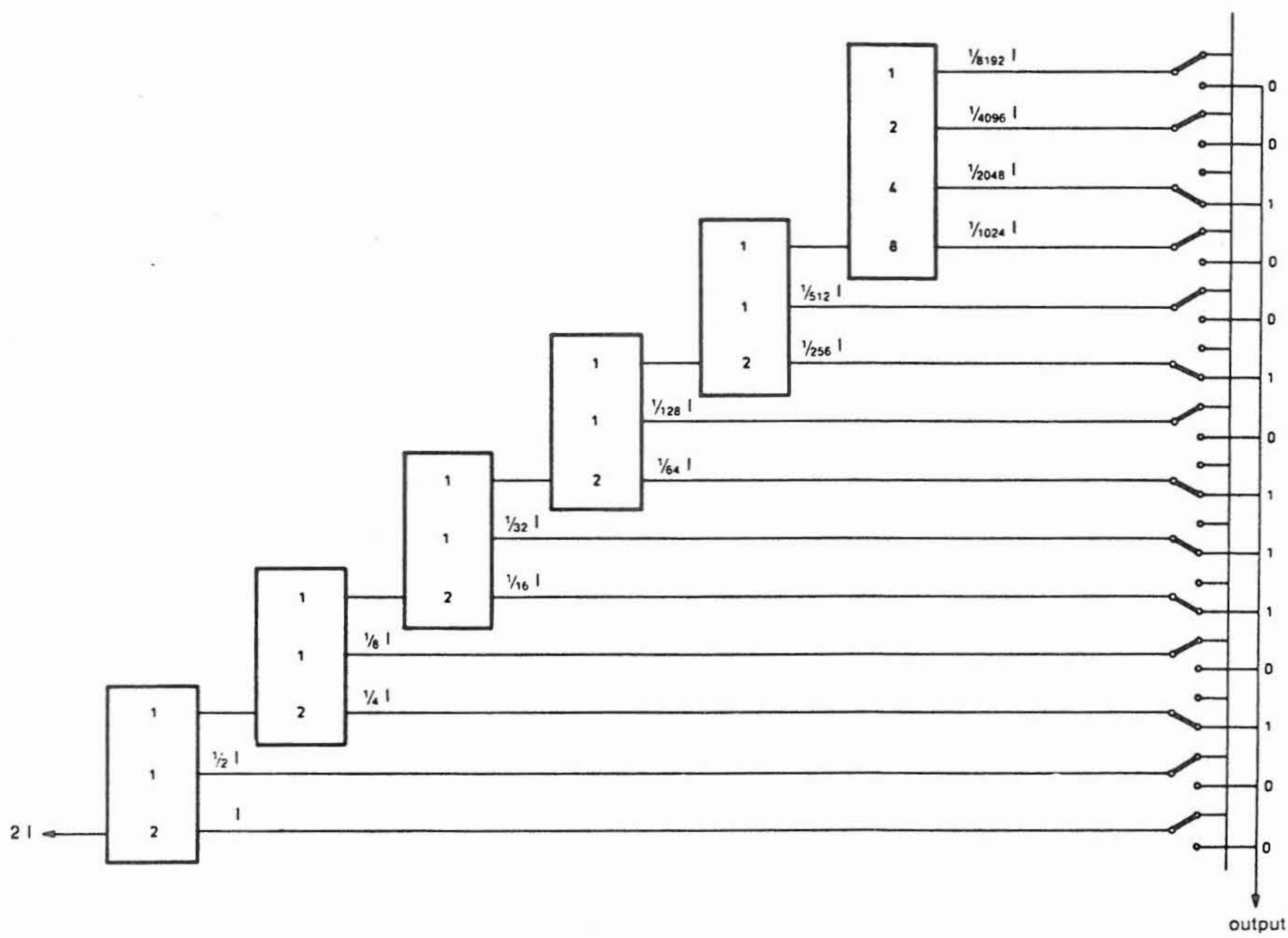
When I = 1 mA, the current will then be 0.3632812 mA.

If the code 00101110100101 (2981) had been presented, the current would be

$$\frac{2981}{8192} I = 0.3638916 \text{ mA}$$

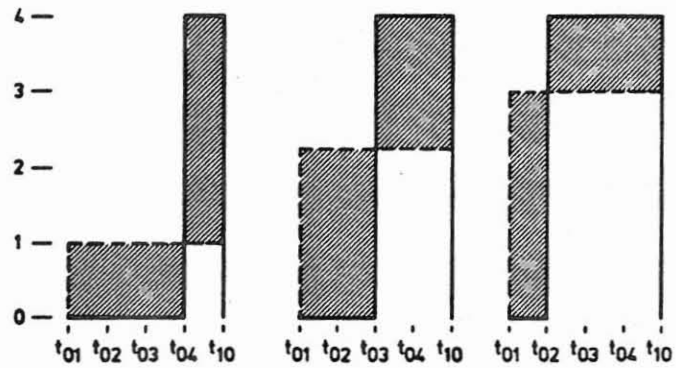
The least significant bit (LSB) results in a change of 0.6104 μA.


LSB
MSB
2 I
I
½ I
¼ I
⅛ I
1/16 I
1/32 I
1/64 I
1/128 I
1/256 I
1/512 I
1/1024 I
1/2048 I
1/4096 I
1/8192 I
2^{0} mA
2^{-13} mA


2 I
I
½ I
¼ I
⅛ I
1/16 I
1/32 I
1/64 I
1/128 I
1/256 I
1/512 I
1/1024 I
1/2048 I
1/4096 I
1/8192 I
output

How can a 14 bit DAC with our times oversampling perform like a 16 bit DAC.

A 14-bit DAC can only output 2^{14} levels (16384).
A 16-bit DAC can output 2^{16} levels (65536).
This means that a 16 bit DAC can output the levels 0 up to 4, while a 14 bit DAC only can output the levels 0 or 4.
Level 1 however can be obtained by keeping the output 0 from t_{01}, up to t_{04} and making the output high from t_{04} to t_{10}, the average output during t_{01} to t_{10} is than equal to level 1.



To obtain an average level 2 the output is low up to t_{03} and high (level 4) from t_{03} up to t_{10}.
Level 3 is obtained by keeping the output low only from t_{01} to t_{02}.