# INTRODUCTION

## ■ The Voice Synthesis LSI

In regular tape recorders where sound is stored as analog signals on magnetic tape, the subsequent amplification of those signals to drive a speaker is referred to as sound "reproduction". It cannot be called sound "synthesis". If the amount of information stored (which is 100% in the above tape recorder) is reduced, the amount of storage medium (memory) can also be reduced for greater economy. Where the amount of information is reduced (compressed) on the basis of a certain principle, the restoration of the original sound is called "synthesis".

## ■ Voice Synthesis Methods

Voice synthesis methods can be divided into three major types:- waveform encoding, parametric synthesis, and synthesis by rule. Each type is briefly outlined below.

(1) Waveform coding methods
This type includes differential PCM (DPCM), adaptive delta modulation (ADM), and adaptive differential PCM (ADPCM). The original sound wave amplitude is sampled at fixed intervals, digitized, and the volume of data then reduced on the basis of the principles mentioned above.

(2) Parametric synthesis methods
Characteristic information included in voice waveforms is extracted as parameters for synthesizing purposes. The PARCOR method is a typical example.
In this method where models of the human vocalization mechanism are used, voiced and voiceless consonant sounds are discriminated, and voiced sound pitch and amplitude data is extracted together with filter characteristic of vocal tract. Voice synthesis is then achieved by passing this data to hardware consisting of digital filter circuits etc.

(3) Synthesis by rule method
This synthesis method is an ideal method where groups of phonemes expressed by small quantities of data are skillfully linked together to reproduce any desired words.
However, since further elucidation of linguistic laws taking intonation, accents, and length sounds into consideration is required to achieve a highly natural voice, this method must still be considered to be in the research stage.

## ■ Basic ADPCM Method

Oki voice synthesis LSIs are based on adaptive differential PCM (ADPCM) which is an improved form of the DPCM method. The PCM and ADPCM methods are described below.

(1) Pulse code modulation (PCM)
Voice waveforms and other analog data can be PCM encoded (into digital data) by sampling and quantization in S&H (sample and hold) and AD converter stages. The S/N ratio in this case is determined by the following expression

$$(N - 1) \times 6 \text{ dB}$$

where N is the number of output bits from the AD converter assuming that the maximum waveform amplitude is equivalent to full scale in the converter. For example, if the sampling frequency is 8 kHz and the number of AD converter output bits is 12, the required amount of data (number of bits) per second is

$$8 \text{ kHz} \times 12 \text{ bits} = 96\text{K bits/sec}$$

This is called the bit rate. The S/N ratio is this case is obtained as

$$(12 - 1) \times 6 = 66 \text{ dB}$$

(2) Adaptive Differential PCM
The ADPCM method was devised as a means of reducing the bit rate without sacrificing the S/N ratio too much. In this method, the amount of data is reduced by quantizing and encoding the differential (dn) between adjacent signal samples. A feature of this method is the ability to make adaptive changes to the quantization width $\triangle n$ when quantizing the differential dn. (In the DPCM method, the quantization width is fixed.) In other words, $\triangle n$ is enlarged when the differential dn is large, and reduced when dn is small.

(3) ADPCM analysis
If the input at the nth sampling point is Xn, and the waveform reproduction value at the (n-1)th sampling point is $\hat{X}n\text{-}1$, the differential dn between the two is

$$dn = Xn - \hat{X}n\text{-}1 \qquad \text{(Differential calculation)}$$

This is then encoded by the quantization width $\triangle n$ at the present point of time (the nth point of time).

$$Ln = dn/\triangle n \qquad \text{(Encoding: Ln = ADPCM data)}$$

And if this is then quantized and the waveform subsequently reproduced,

$$qn = (Ln + 1/2)\triangle n \qquad \text{(Quantization)}$$

$$\hat{X}n = \hat{X}n\text{-}1 + qn \qquad \text{(Reproduction)}$$

The quantization width for the next (n+1)th item of data is then changed from $\triangle n$ to $\triangle n+1$.

$\Delta n+1 = \Delta n \cdot M(Ln)$ (Quantization width change)
(where M(Ln) is the Ln function format)

The quantization width is thus determined according to the previously accumulated data.

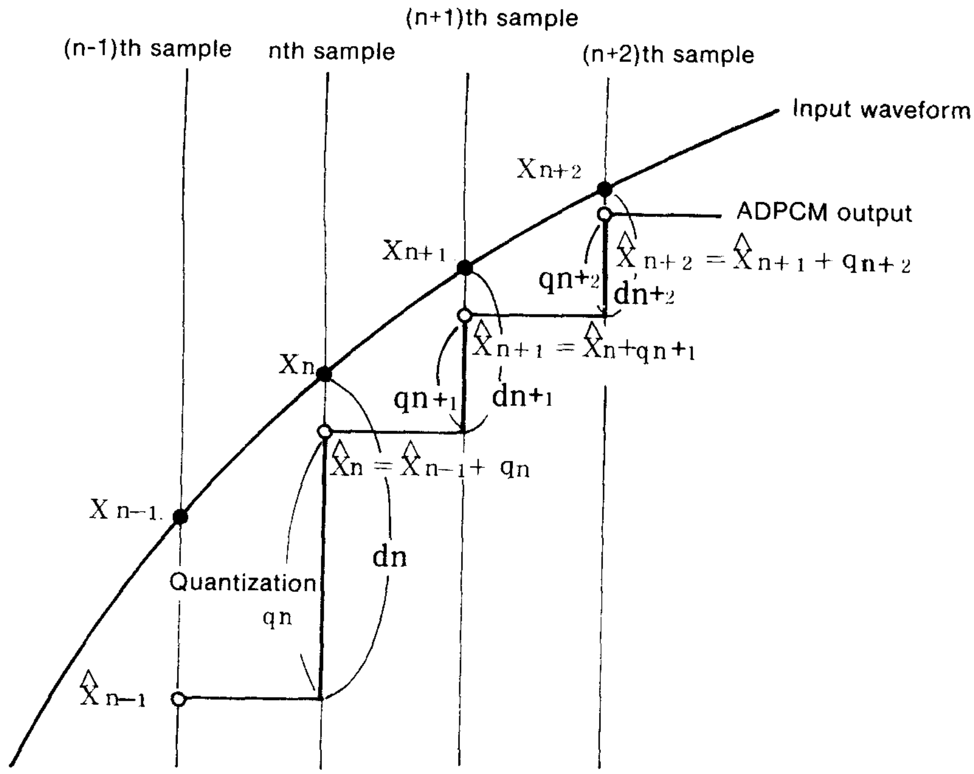



**Fig. 1 ADPCM method computation process**

(4) ADPCM data
The ADPCM data Ln is expressed as M bits of data including the polarity bit. For example, a 4-bit expression is given as

$$Ln = \{B_3, B_2, B_1, B_0\}$$

where $B_3$ (polarity bit) indicates the polarity of the (begin-ind)differential dn,
$B_2$ (MSB) indicates the presence of a $4\Delta n$ digit in the changed portion,
$B_1$ (2SB) indicates the presence of a $2\Delta n$ digit in the changed portion, and
$B_0$ (LSB) indicates the presence of a $\Delta n$ digit in the changed portion.

(5) ADPCM reproduction
ADPCM reproduction can be expressed as part of the ADPCM analysis. The reproduction for the nth ADPCM input data Ln is given as

$$qn = (Ln + 1/2)\triangle n \quad \text{(Quantization)}$$
$$= (-1 \times B_3)\ (4\triangle nB_2 + 2\triangle nB_1 + \triangle nB_0 + \tfrac{1}{2}\triangle n)$$
$$\hat{X}n = \hat{X}n\text{-}1 + qn \quad \text{(Reproduction)}$$

And $\triangle n+1$ is calculated for the (n+1)th item of ADPCM data.

$$\triangle n+1 = \triangle n \cdot M(Ln)$$

In other words, in addition to serving as data used for calculating new PCM values for previously set quantization widths, the ADPCM code also serves as the data for calculating the quantization width to be set next. Furthermore, if

Xn is set to 12 bits, and
Ln is set to 4 bits,

the quantity of ADPCM encoded data is compressed by 4/12.

Note: The 1/2 element in
quantization $qn = (Ln + 1/2)\triangle n$
serves as a means of linear equalization for changed polarity.

## ■ Oki ADPCM Types

Although the ADPCM method is the basic method adopted by Oki, a few modifications have been made, and two analysis/synthesis methods are now in use. These are outline briefly below.

(1) Straight ADPCM: This is the basic unmodified ADPCM method where the quality of sound is better than the two methods described below. This method is also suitable for sound effects.

(2) Compressed ADPCM: The compressed ADPCM method is the straight ADPCM method subject to unvoice elimination processing[Note 1] and waveform repetition processing[Note 2]. The bit rate, therefore, can be reduced to 1/3 of the straight ADPCM bit rate. Furthermore, the degree of data compression can be changed for each word.

Note 1. Unvoice elimination processing:
Extensive reduction of unvoice interval data by replacing unvoice intervals (which exceed a certain length) with unvoice data.

Note 2. Waveform repetition processing:
Data volume reduced by repeating a single item of waveform data in voiced waveforms such as vowels which are repeated periodically.

## ■ Features of Oki Voice Synthesis LSIs

The major features of Oki voice synthesis LSIs are summarized below.

(1) Quality and object of synthesis
- (i) Good quality sound with high degree of naturalness
- (ii) Synthesis of sound effects, musical instruments, and animal sounds also possible

(2) Hardware
- (i) Easy to handle built-in ROM 1-chip devices prepared for application in simple sets
- (ii) Range of voice synthesis LSIs with varying built-in ROM sizes to meet diversified market needs
- (iii) Low power requirements due to CMOS with low fundamental oscillating frequency — ideal also for battery operated applications

(3) Software
- (i) Simple and precise analysis for broader range of user selected sounds
- (ii) Comprehensive range of analytical tools to enable synthesis to be executed by user

Because of the fine quality of sound achieved in a wide range of applications, Oki voice synthesis LSIs are used by a great many users in many different applications.

# PRODUCT LINE-UP

| Functions and Specifications \ Device | | 5205 | 5218 | 6295 | 5248 | 6243 | 6212 | 6258 Stand-alone | 6258 MPU interface | 6308 | 6309 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ADPCM Method | Straight | ○ | ○ | ○ | ○ | ○ | ○ | ○ | | ○ | ○ |
| | Compressed | — | — | — | — | ○ | ○ | — | | — | — |
| ADPCM Bit Length | 3 bit | ○ | ○ | — | ○ | — | ○ | ○ | | — | — |
| | 4 bit | ○ | ○ | ○ | — | ○ | — | ○ | | ○ | ○ |
| Oscillation Frequency (Hz) | | 384 ~ 738K | 384 ~ 768M | 1 ~ 5M | 10 ~ 50K | 30 ~ 132K | 30 ~ 132K | 4 ~ 8 M | | 4 ~ 6 M | 4 ~ 6 M |
| Maximum Vocalization Time | | Determined by external connections | Deternubed by external connections | *1 90 sec. | *1 3 sec. | *2 20 sec. | *2 40 sec. | Determined external connections | | Determined external connections | Determined external connections |
| Maximum Number of Vocalized Words | | Determined by external connections | Determined by external connections | 127 | 7 | 124 | 124 | 7 | Deter-mined by ex-ternal connec-tions | 4 | 4 |
| Analytical Functions | | — | ○ | — | — | — | — | ○ | | ○ | ○ |
| Built-in ROM (Kbits) | | — | — | — | 48 | 192 | 288 | — | | — | — |
| Built-in DA Converter | | Voltage type | Voltage type | Voltage type | Voltage type | Current type | Voltage type | Voltage type | | Voltage type | Voltage type |
| Built-in AD Converter | | — | — | — | — | — | — | ○ | | ○ | ○ |
| External Memory Interface | | — | — | ○ | — | — | — | ○ | | ○ | ○ |
| Input Interface | SW Input | — | — | — | ○ | ○ | ○ | ○ | — | ○ | ○ |
| | Microcomputer Control | ○ | ○ | ○ | ○ | ○ | ○ | ○ | | ○ | ○ |
| DA Output | A class | ○ | ○ | ○ | — | ○ | ○ | ○ | | ○ | ○ |
| | B class | — | — | — | [illegible] | ○ | ○ | — | | — | — |
| MSC1161 Connection | | — | — | — | — | ○ | ○ | — | | — | — |
| Power Supply Voltage (V) | | 5 | 5 | 5 | 3 | 3 or 5 | 3 or 5 | 5 | | 5 | 5 |
| Power Consumption (in standby mode) | | 4 mA (−) | 4 mA (−) | 5 mA (−) | 0.5 mA (1.0 μA) (0.5μA) | Note 3 0.5 mA (0.5μA) | Note 3 0.5 mA | 4 mA (−) | | 10 mA (−) | 10 mA (1.0μA) |
| Package Shape | | 18 DIP | 24 DIP 32 flat | 44 flat | 18-DIP 24 flat chip | 40-DIP 44, 60 flat chip | 40-DIP 44, 60 flat chip | 60 flat | 40-DIP 44 flat | 44 flat | 60 flat |
| Remarks | | — | — | External ROM | ROM code device | ROM code device | ROM code device | — | | — | — |

Note: *1 Applicable when $f_{SAMPLE} = 5.5$ kHz
*2 Applicable when $f_{SAMPLE} = 8.2$ kHz compressed ADPCM
*3 Applicable when power supply voltage is 3 V

**OKI semiconductor**

# MSM5205

## ADPCM SPEECH SYNTHESIS IC

## GENERAL DESCRIPTION

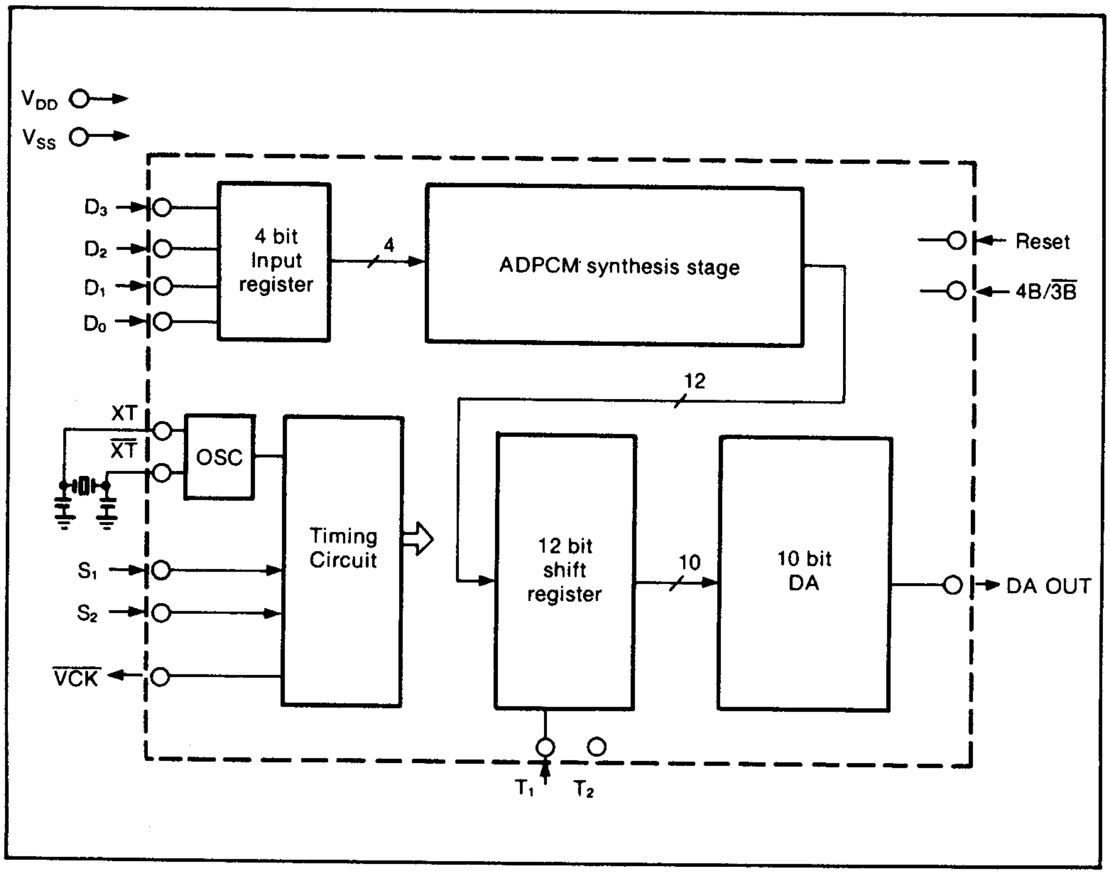
The MSM5205 is a speech synthesis integrated circuit which accepts Adaptive Differential Pulse Code Modulation (ADPCM) data. The circuit consists of synthesis stage which expands the 3- or 4-bit ADPCM data to 12-bit Pulse Code Modulation (PCM) data and a D/A stage which reproduces analog signals from the PCM data.

The MSM5205 is fabricated using Oki's advanced CMOS process which enables low power consumption. The single power supply requirement and its availability in 18-pin molded DIP allow the MSM5205 to be ideally suited for various applications.

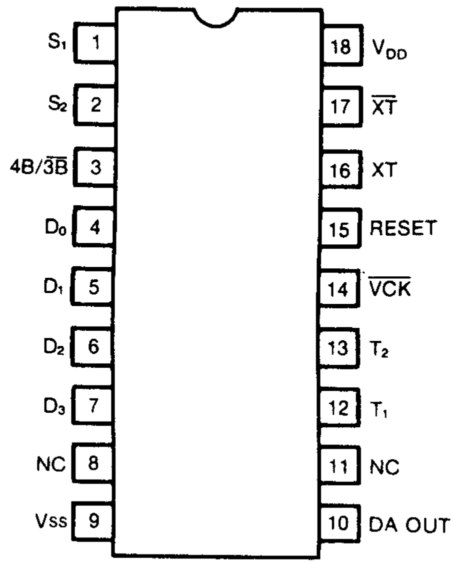
## FEATURES

- 3 or 4 bit ADPCM system
- 12 to 32 kb/sec with INT $\overline{VCK}$
- On-chip 10-bit D/A converter
- Low power consumption (10 mW typical)
- Single +5V supply
- Wide operating temperature (Ta = -30° C to +70° C)
- 18-pin molded DIP

## FUNCTIONAL BLOCK DIAGRAM



## PIN CONFIGURATION




## ABSOLUTE MAXIMUM RATINGS

| Parameter | Symbol | Conditions | Ratings | Unit |
|---|---|---|---|---|
| Power supply voltage | $V_{DD}$ | Ta = 25° C | -0.3 to +7.0 | V |
| Input voltae | $V_{IN}$ | Ta = 25° C | -0.3 to $V_{DD}$ | V |
| Power dissipation | $P_D$ | Ta = 25° C | 200 max | mW |
| Storage temperature | Tstg | — | -55 to +150 | °C |

**Note:** Stresses above those listed under ABSOLUTE MAXIMUM RATINGS may cause permanent damage to the device. This is a stress rating only and functional operation of the device at these or at any other condition above those indicated in the operational section of this specification is not implied. Exposure to absolute maximum rating conditions for extended periods may affect device raliability.

## OPERATING CONDITIONS

| Parameter | Symbol | Conditions | Ratings | Unit |
|---|---|---|---|---|
| Power supply voltage | $V_{DD}$ | — | +3 to +6 | V |
| Operating temperature | Top | — | -30 to +70 | °C |

## D.C./A.C. CHARACTERISTICS

($V_{DD}$ = 5V ± 5%; Ta = -30° C to +70° C, unless otherwise noted)

| Parameter | Symbol | Conditions | Min. | Typ. | Max. | Unit |
|---|---|---|---|---|---|---|
| Input High Voltage | $V_{IH}$ | All inputs except $T_1$, $T_2$ | 4.2 | — | $V_{DD}$+.3 | V |
| Input Low Voltage | $V_{IL}$ | All inputs except $T_1$, $T_2$ | $V_{SS}$—.3 | — | 0.8 | V |
| Input High Current | $I_{IH}$ | $V_{IN} = V_{DD}$ | — | — | 1 | $\mu$A |
| Input Low Current | $I_{IL}$ | $V_{IN}$ = 0V | — | — | -1 | $\mu$A |
| Output High Current | $I_{OH}$ | $\overline{VCK}$ pin: Vo = 4.2V | -50 | — | — | $\mu$A |
| Output High Current | $I_{OL}$ | $\overline{VCK}$ pin: Vo = 0.4V | +50 | — | — | $\mu$A |
| Oscillator Frequency | $f_{OSC}$ | Specified Oscillator | — | 384 | 768 | kHz |
| Operating Current | $I_{DD}$ | $f_{OSC}$ = 384 kHz<br>$V_{DD}$ = 5V | — | 2 | 4 | mA |
| D/A Accuracy (Internal 10-bit D/A) | $V_E$ | Full Scale; $V_{DD}$ = 5V | — | ±4 | — | LSB |
| $DA_{OUT}$ Output Impedance | $V_{OR}$ | | — | 100 | — | KΩ |

## PIN DESCRIPTION

| Pin Name | Terminal Number | I/O |
|---|---|---|
| $S_1$ | 1 | I |
| $S_2$ | 2 | I |

These inputs select the sampling frequency according to Figure 1.

| | | |
|---|---|---|
| $4B/\overline{3B}$ | 3 | O |

Specifies whether 3-bit or 4-bit ADPCM data is to be processed.

| | | |
|---|---|---|
| $D_0$ | 4 | I |
| $D_1$ | 5 | I |
| $D_2$ | 6 | I |
| $D_3$ | 7 | I |

ADPCM data inputs. For 3-bit ADPCM data, $D_0$ input is not used and should be connected to ground.

| | | |
|---|---|---|
| $V_{SS}$ | 9 | I |

Ground (0 V)

| | | |
|---|---|---|
| $DA_{OUT}$ | 10 | O |

Output for synthesized analog signal. Peak-to-peak swing is proportional to $V_{DD}$. Typical connection scheme is shown Figure 2.

## PIN DESCRIPTION (continued)

| | | |
|---|---|---|
| $T_1$ | 12 | I |
| $T_2$ | 13 | I |

IC test pins used at the factory for testing purposes only. During normal operations, $T_1$ is grounded and $T_2$ is left open.

| Pin Name | Terminal Number | I/O |
|---|---|---|
| $\overline{VCK}$ | 14 | O |

This pin outputs a signal whose frequency is equal to the sampling frequency selected by the $S_1$, $S_2$ inputs. See note *1.

| | | |
|---|---|---|
| RESET | 15 | I |

An active high input which initializes the internal circuitry. Internally, the reset pulse is synchronized with the $\overline{VCK}$ signal. To be effective, it must be true for at least twice $\overline{VCK}$ time.

| | | |
|---|---|---|
| XT | 16 | I/O |
| $\overline{XT}$ | 17 | I/O |

Oscillator input and output for a 384 kHz crystal or ceramic resonator (Figure 3).

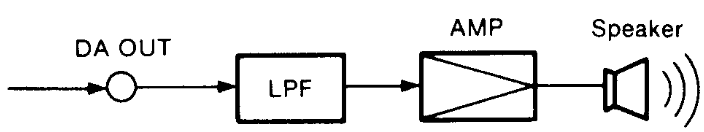
| | | |
|---|---|---|
| $V_{DD}$ | 18 | I |

Power supply pin (Typical +5 V)

| S1 | S2 | Sampling Frequency |
|---|---|---|
| L | L | 4 kHz (384 kHz/96) |
| L | H | 6 kHz (384 kHz/64) |
| H | L | 8 kHz (384 kHz/48) |
| H | H | Prohibited See Note *1 |

Note: *1 The 384 kHz oscillator must be used whether 4 kHz, 6 kHz, 8 kHz.

**Figure 1 Functional table**




Cut off frequency fc of LPF should be related to the selected sampling frequency f sample by,

$$fc = f\ sample/2 \times 0.85$$

Sound quality is strongly dependent on the sharpness of the low pass filter.

*If the 5205 is sent a stream of ADPCM data that causes greater than full scale output, the D/A output will wrap around: +5 0.0 +5.

**Figure 2**

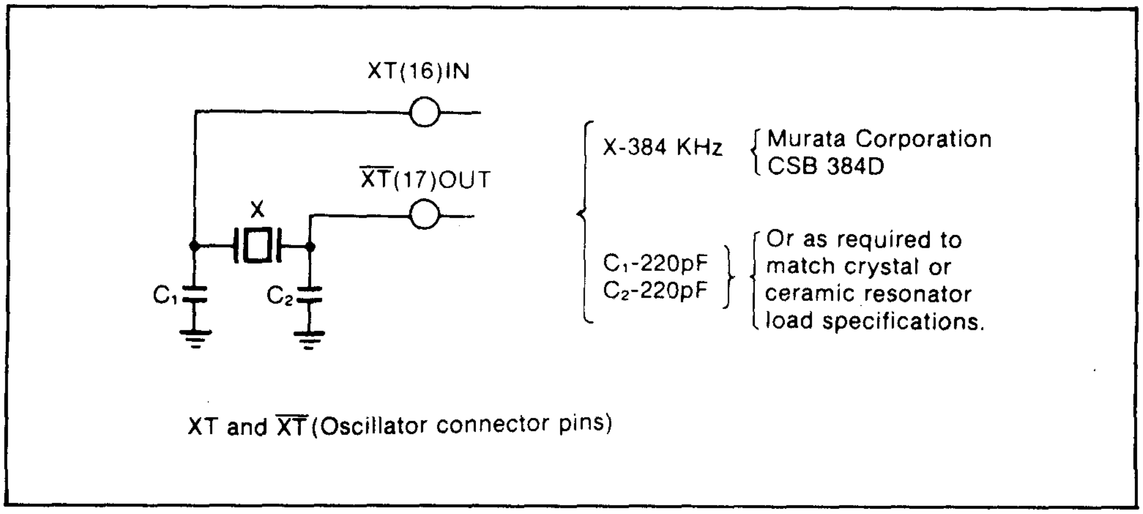



**Figure 3**

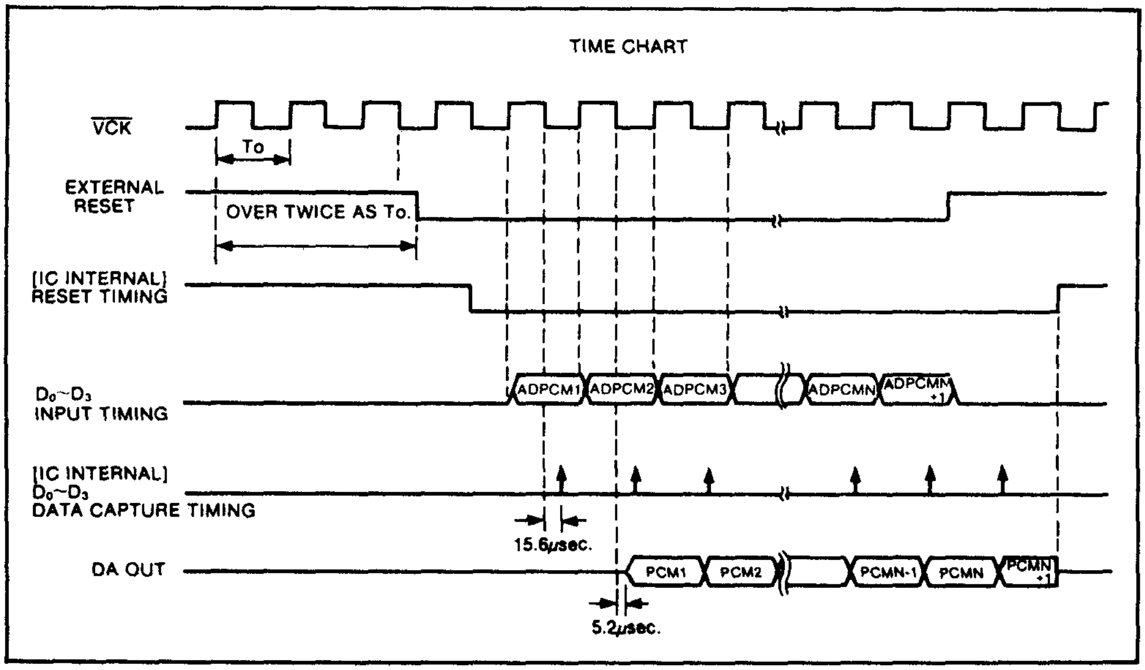



**Figure 4**

## DISTINCTION BETWEEN MSM 5218 AND MSM5205

Both Synthesis stages (MSM5218 and MSM5205) work with the same method. However, with the exception that MSM5218 is equipped with an overflow protection.
In other words, when all 12 PCM bits become '1' any further exceeding analog input would cause a data overflow which is catched and re-routed as the MSB in case of MSM5218.
MSM5205 returns to 'all bits zero' when a data overflow sets in.
Therefore, the DA output of MSM5205 is distorted badly.

When MSM5218 is being used to generate ADPCM data for playback on MSM5205, the peak to peak input level to the A to D converter should be limited to 80% of the converters maximum input range. The use of an automatic gain control (AGC) amplifier or a hard limiter is recommended.

## TYPICAL APPLICATION

**MSM5205 to Centronics Interface Circuits** ($f_{SAMPLE}$ = 8kHz)

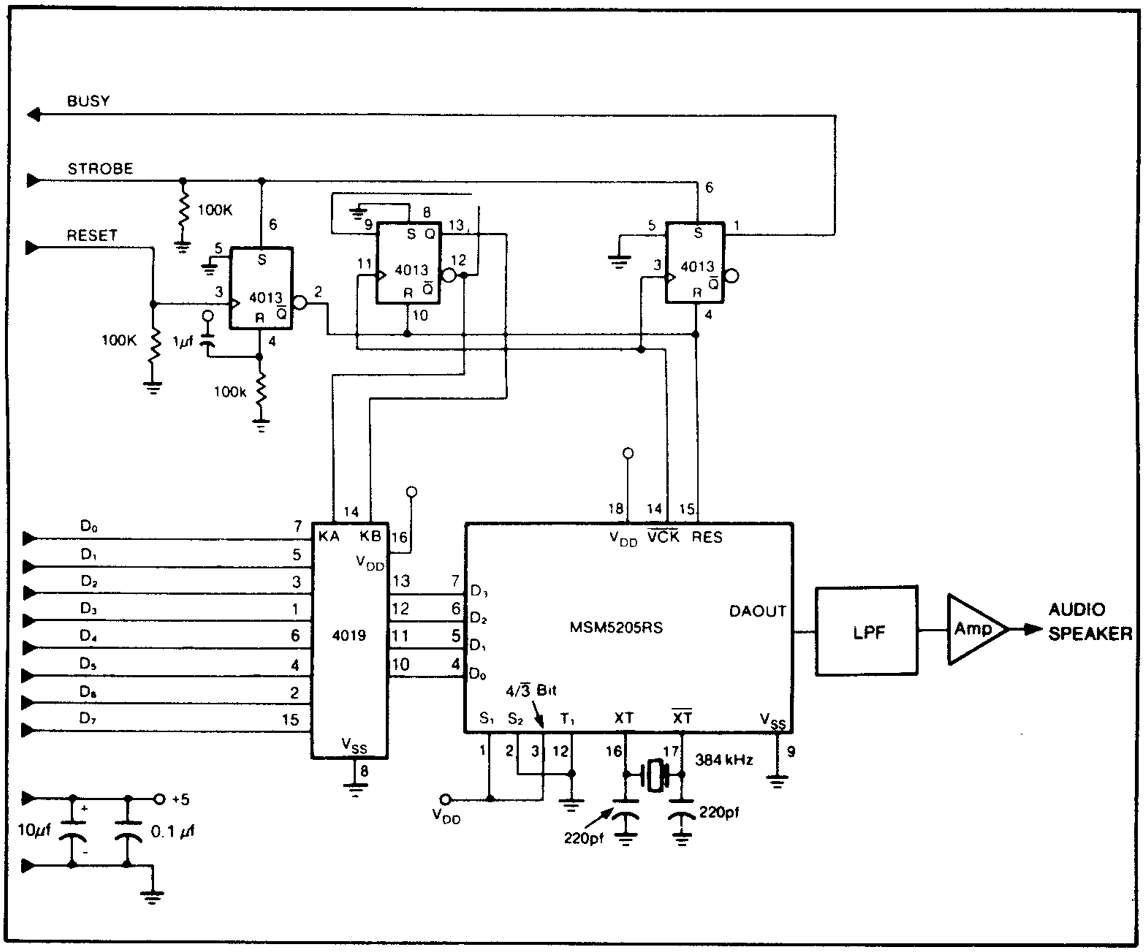



**Figure 5**

**MSM5205 to Centronics Timing Diagram**

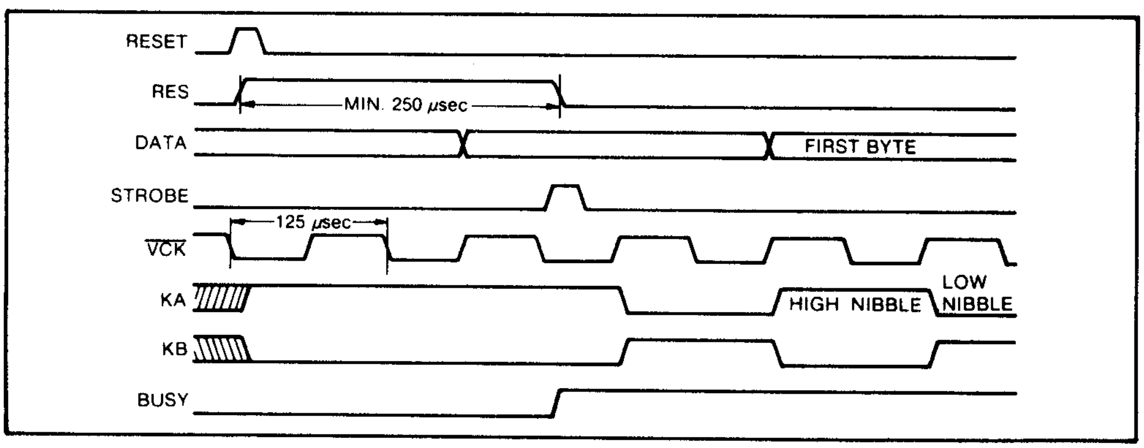



**Figure 6**

# Q & A

**Q1** What is the relationship between the vocalization time and quality in built-in ROM voice synthesis LSIs ?

**A1** Parameters involved in determining vocalization time and quality include the sampling frequency, ADPCM word length, and the degree of data reduction when a compressed ADPCM method is used. These parameters are described below.

(1) Sampling frequency:
The vocalization bandwidth where voice can be synthesized is determined by the sampling frequency. According to the sampling theorem,

$$f_{BAND} = f_{SAMPLE} \times 1/2$$

where $f_{BAND}$ is the upper limit of the frequency passband and $f_{SAMPLE}$ is the sampling frequency
The effective bandwidth ($f_{BAND}$ eff) where the ADPCM method response (see Figures 33 and 34) and the LPF used are considered can be stated as:

$$f_{BAND}\ \mathrm{eff} = f_{SAMPLE} \times 1/2 \times 0.85$$

The hearing evaluation for different $f_{SAMPLE}$ values are listed in Table 1.

**Table 1**

| $f_{SAMPLE}$ | Bandwidth | Hearing evaluation |
|---|---|---|
| 8.2 kHz | DC ~ 3.49 kHz | Very clear sound including almost all vocalized sounds |
| 6.55 kHz | DC ~ 2.78 kHz | High tone female voices sound as if the nose is blocked up |
| 4.1 kHz | DC ~ 1.78 kHz | Both male and female voices sound as if the nose is blocked up |

The "blocked up nose" sound is due to elimination of high frequency components normally included in the human voice. Since high frequency components are prevalent in constants (and in the "i" vowel), the effect is stronger at lower sampling frequencies.
In determining the sampling frequency, it is better to decide the most suitable value after listing to the source voice passed through a LPF.

(2) ADPCM word length:
The ADPCM word length is related to the S/N ratio. The 4-bit ADPCM method is better than the 3-bit ADPCM method by about 6 dB.

(3) Degree of data reduction in compressed ADPCM methods:
The quality of sound varies according to the degree of compression in compressed ADPCM methods. In compressed ADPCM, data can be compressed to 1/3 of the data used in straight ADPCM. Needless to say, the greater the degree of compression, the poorer the quality of sound becomes. The degree of compression is selected according to the intended purpose.

(4) Vocalization duration and quality of sound
An example of the relationship between vocalization duration and the quality of sound as determined by the above parameters when MSM6243 is used is outlined in Figure 29.

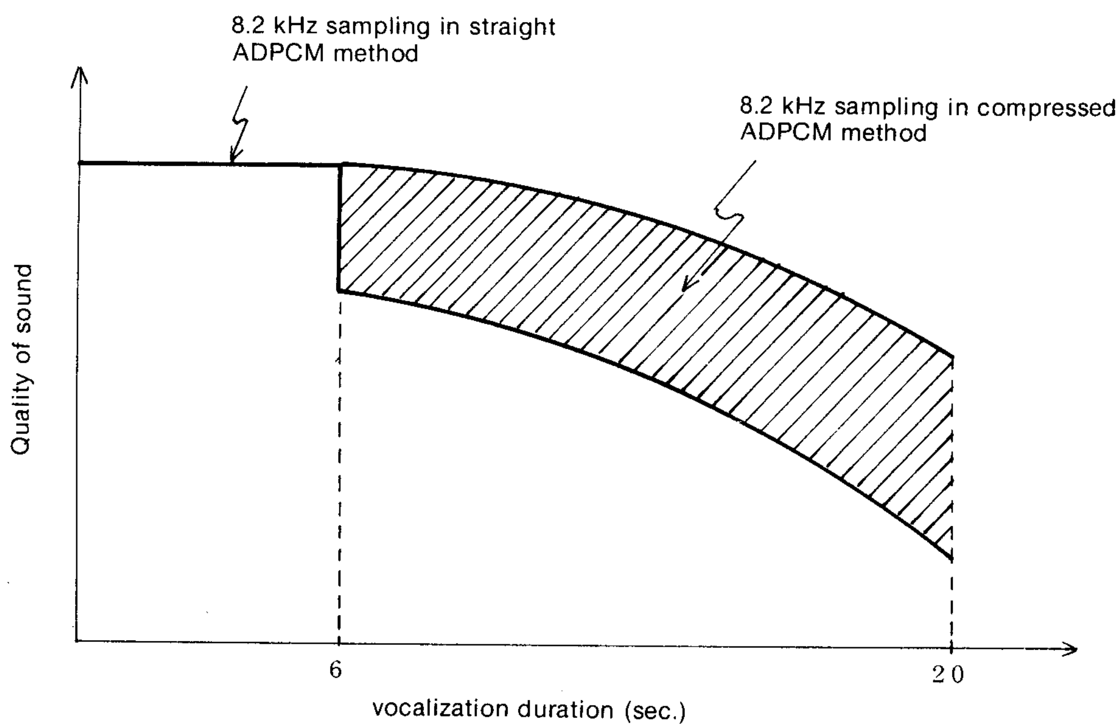



**Fig. 29 Relationship between vocalization duration and quality of sound**

The quality of sound is depicted as a range in the compressed ADPCM method in Figure 29. This is because the quality of sound is also dependent on the quality of the source sound vocalized statement specifications. And since the quality of sound cannot be expressed in absolute terms, it is shown here in relative terms.

(5) Examples of calculating the vocalization duration:
An example of calculation of the maximum vocalization duration in straight ADPCM for MSM6243 at a sampling frequency of 8.2 kHz is given below.

$$193{,}584\ [\text{bits}] \div (\underset{\substack{\uparrow\\ f_{SAMPLE}}}{8.2\ [\text{kHz}]} \times \underset{\substack{\uparrow\\ \text{ADPCM word length}}}{4\ [\text{bits}]}) = 5.90\ [\text{sec.}]$$

↑ Built-in ROM capacity (193,584 [bits])

The reason why the built-in ROM capacity is not 196,608 bits = 192 kbits x 1,024 bits/kbits is because the user will have a certain amount of area where use is disabled.

**Q2** Although there is a certàin amount of variation in the precision of the MSM5205 and MSM5218 DA converters, are there any methods for improving the S/N ratio?

**A2** Yes. In terms of configuration in the above two DA converters, it is possible for the voice waveform precision to fall in the vicinity of the center waveform. Therefore, the S/N ratio can be improved by displacing the center of the waveform either up or down. This method is particularly effective in improving the S/N ratio for low level signals and in minimizing the residual noise during silent periods (such as the intervals between successive words). In practice, the waveform center can be shifted by adding additional data before or after the current ADPCM data (voice data) as indicated in Figure 30.

| Section (A) | Section (B) | |
|---|---|---|
| 0 0 0 0 | 1 0 0 0 | 100 items of data |
| 0 0 0 0 | 1 0 0 0 | |
| ⋮ | ⋮ | |

(Example where the ADPCM bit length is 4 bits)

Since an offset of about 5 mV per 2 items of data can be obtained in this case, shifts case, shifts of about 250 mV will require input of some 100 items of data. The shift is upwards in (A), and downwards in (B). The number of data items should be equal in (A) and (B). The output waveform obtained when (A) is added at the start of the voice data and (B) is added at the end is shown in Figure 30.

Note that in since the dynamic range is compressed in the shifted sections, overflow may occur in some cases depending on the data, and this can result in clipping of the output sound. Reduce the sound pressure by about 20% and repeat the analysis. (See Q3 for additional information about overflow.)

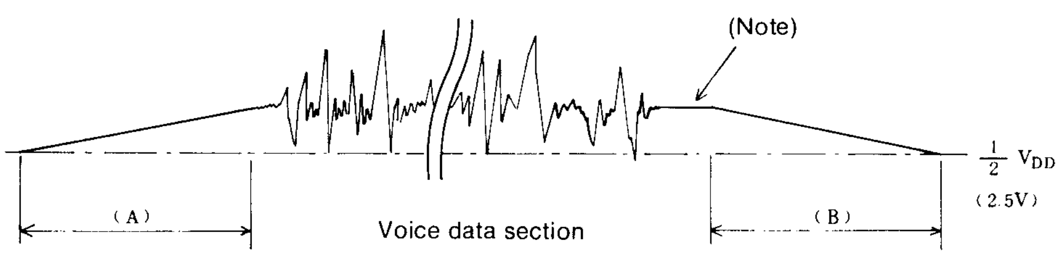



**Fig. 30 DA converter output waveform**

Note: There must be less voice data immediately prior to section (B). Input of some 20 to 30 msec of silence prior to the beginning of the voice is recommended.

**Q3** When ADPCM data analyzed by MSM5218 was used in voice synthesis by MSM5205, noise was generated in the synthesized voice. How can this be overcome? There was no noise in the output from the MSM5218 DAOUT pin.

**A3** The analysis must be repeated.
This is because MSM5205 is not equipped with the overflow prevention circuit featured in the MSM5218 internal computing circuits. Although sound synthesized by MSM5218 may be normal, overflow can occur during synthesis by MSM5205 resulting in the generation of noise. In this case, the data must be re-analyzed and the ADPCM data regenerated.
An example of a waveform where overflow has occurred, and a method for avoiding this overflow are outlined below.

(1) Waveforms where overflow occurs
Observation of the MSM5205 DA converter output waveform by oscilloscope shows the occurrence of an overflow.
(see Figure 31).

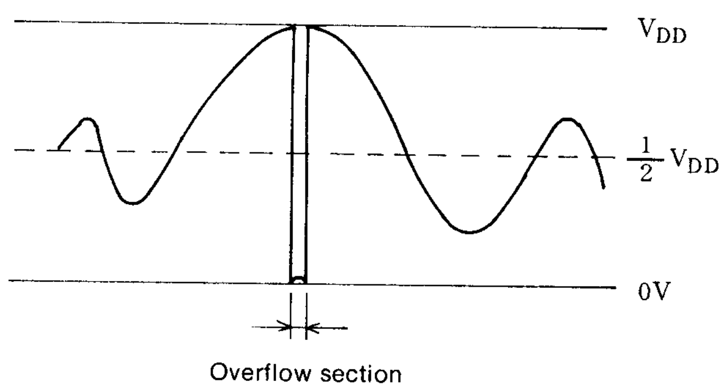



**Fig. 31 Output waveform when overflow occurs**

(2) Method used to avoid overflow
Even where the input waveform does not exceed the dynamic range when analyzed by MSM5218, output overflow can be generated by internal calculation error. Therefore, even if the input amplitude level reaches a maximum when analyzed by MSM5218, keeping the level within 80% of the dynamic range (see Figure 32) will ensure that no overflow occurs in the MSM5205 output and no noise is generated in the synthesized sound.

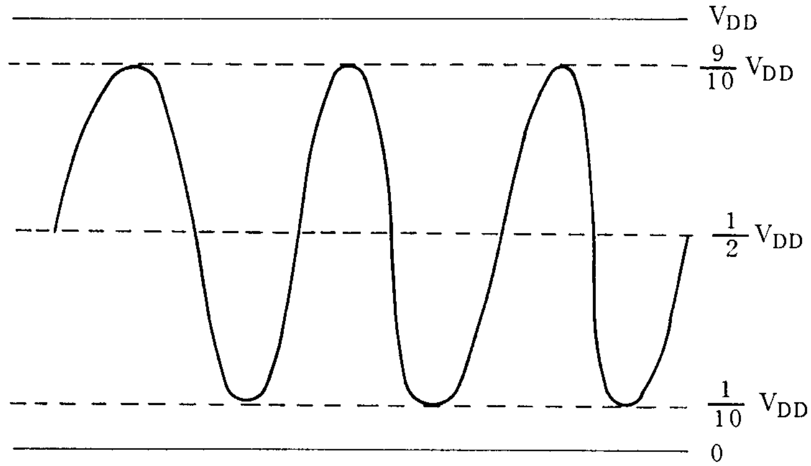



**Fig. 32 Waveform with amplitude kept within 80% of the dynamic range**

**Q4** What kind of frequency response is achieved in the output waveform in respect to Oki ADPCM input waveforms?

**A4** The frequency response achieved in output waveforms when sinewave input waveforms are applied is outlined in Figures 33 and 34.
Since these diagrams show the frequency response when 4-bit ADPCM data is used with an 8 kHz sampling frequency, an increase in the sampling frequency will result in a shift of the frequency response to the right.

(1) When input waveform is a 1/2 scale sinewave (1/2VDD Vp-p)

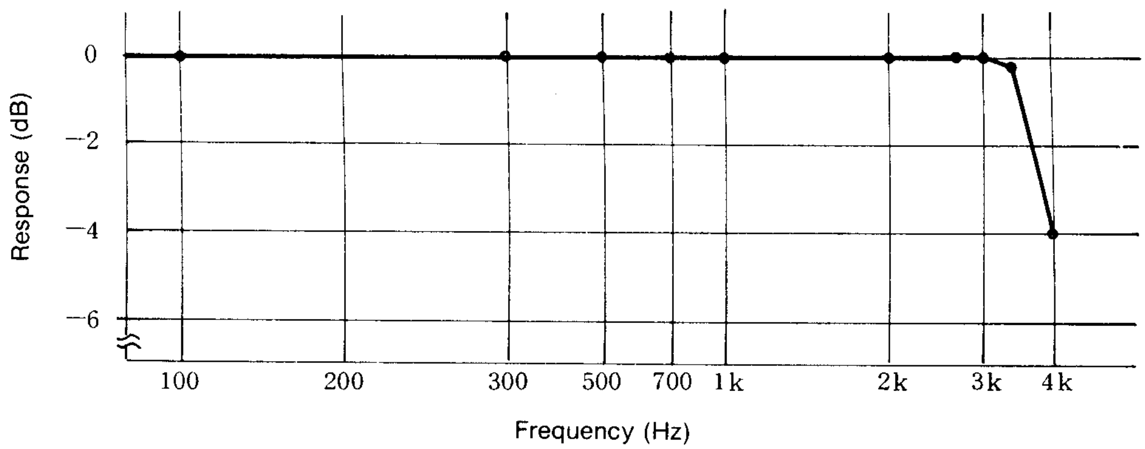



**Fig. 33 Oki ADPCM frequency response (1)**

(2) When input waveform is a full scale sinewave (VDD Vp-p)

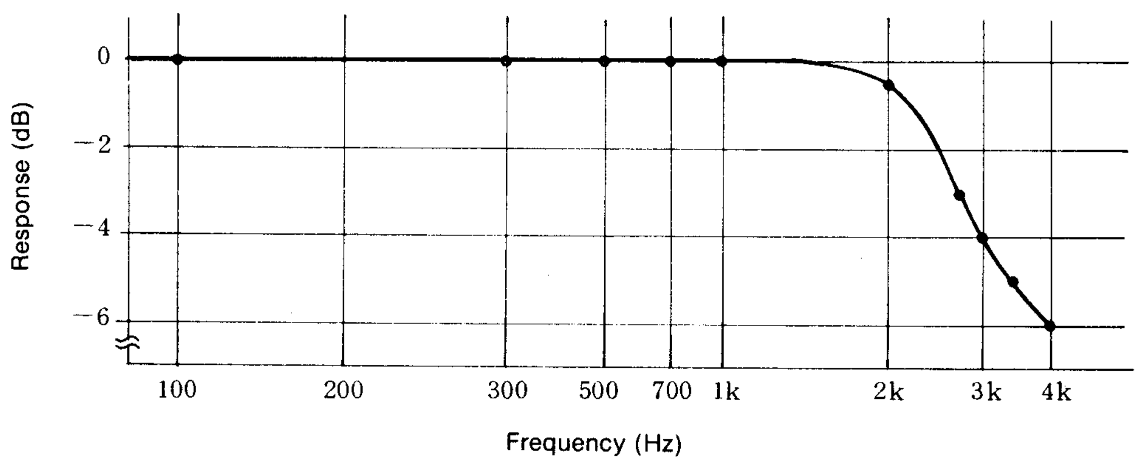



**Fig. 34 Oki ADPCM frequency response (2)**

**Q5** What does it mean to say that the DA converter output from MSM6243 etc is a "class A output" or "class B output"?

**A5** MSM6243 and other devices are equipped with two DA converter output pins DAU and DAL. In class A output mode (the normal mode), the output waveform only appears at the DAU pin with the VDD/2 potential at the amplitude center to obtain maximum amplitude for Vss thru VDD. Class B output mode is used in interfacing with MSC1161 (note). The upper half of the waveform above the class A output amplitude center appears at the DAU pin and the lower half of the waveform appears at the DAL pin with the main amplitude direction as the positive direction. The comparative waveforms are shown in Figure 35.

Where class B is multiplied by 1, the amplitude scale is twice the class A level, and likewise class B multiplied by 2 and 4 result in class A multiplied by 4 and 8 respectively.
In waveforms where the VDD level is exceeded, however, the output is clamped at $V_{DD}$.

Note: The method described in MSM6212 Switch Input Interface Example at page 174.

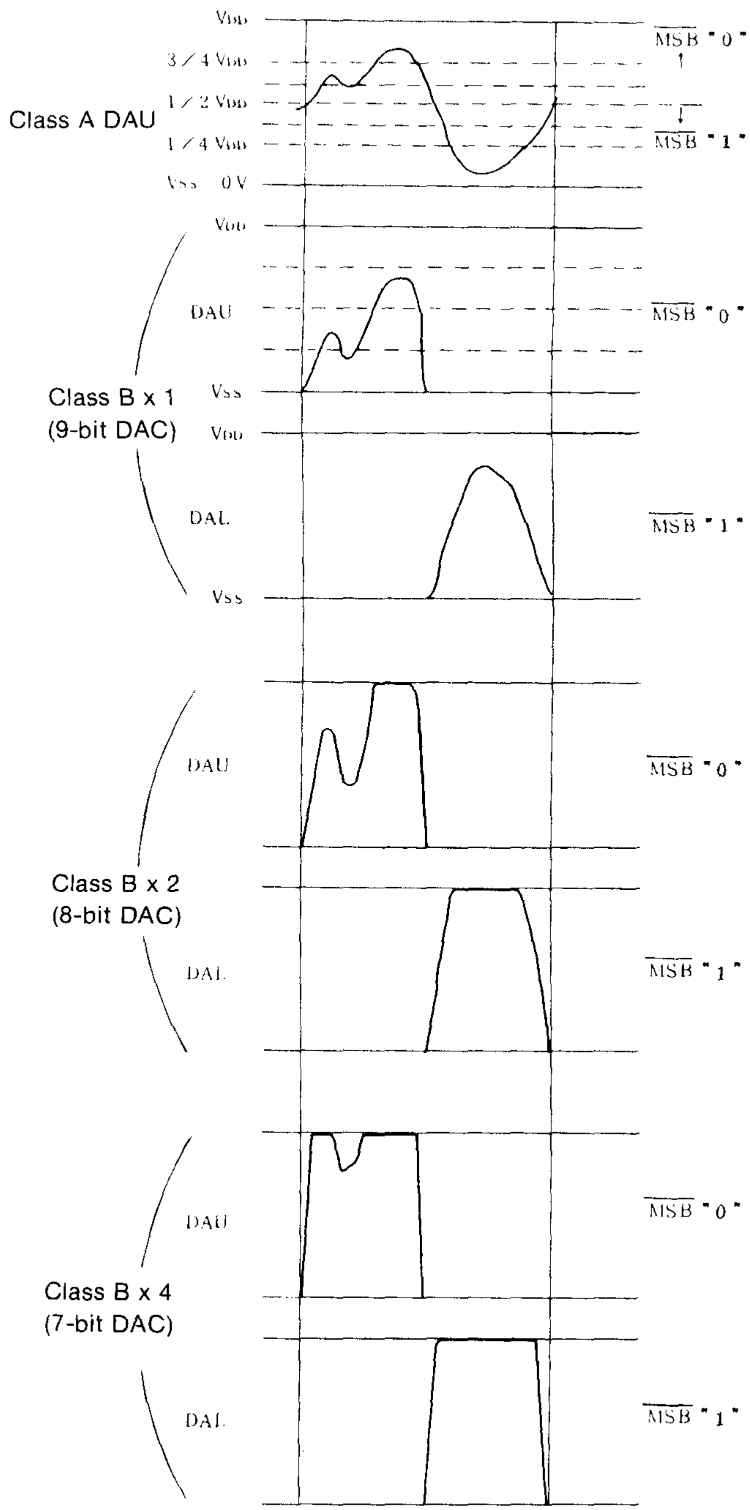



**Fig. 35 Comparative waveforms**

**Q6** What is the best way to form a sinewave with MSM5205?

**A6** By input of the following data values. The corresponding output waveforms are also shown. The data used is 4-bit data (in hexadecimal notation).

(1) To obtain an output of 1.5 Vp-p
Use the following input data.

0,0,0,3,3,3,7,F,7,F,4,C,4,C
Then 4,0,9,C,8,1 repeatedly.

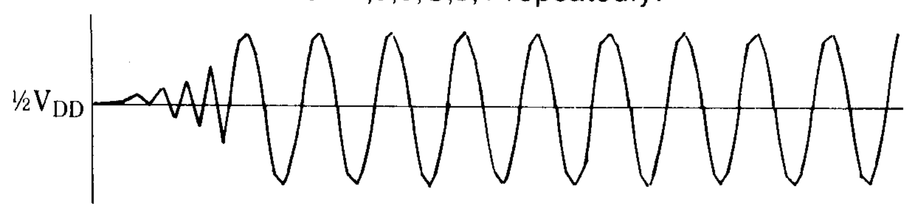



**Fig. 36 1.5 Vp-p output**

(2) To obtain an output of 0.85 Vp-p
Use the following input data.

0,0,0,B,B,B,7,F,7,F,4,C,0,0
Then 4,0,9,C,8,1 repeatedly.

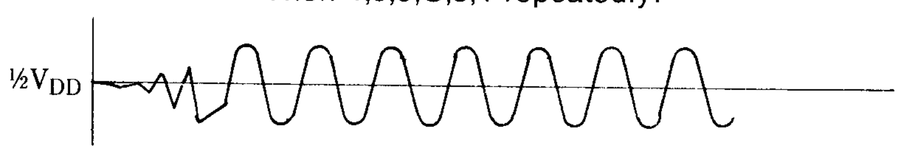



**Fig. 37 0.85 Vp-p output**

(3) To obtain an output of 2.6 Vp-p
Use the following input data.

0,0,0,0,7,F,7,F,7,F,4,C,0,0
Then 4,0,9,C,8,1 repeatedly.

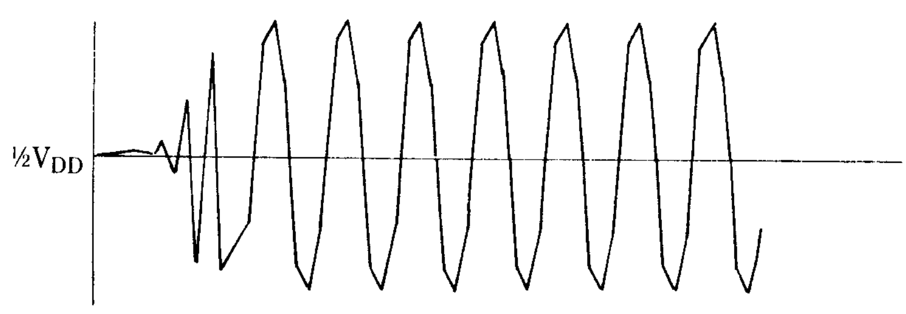



**Fig. 38 2.6 Vp-p output**

**Q7** How should the MSM5205 and MSM5218 reset input timing be best set?

**A7** The MSM5205 and MSM5218 reset should be input according to the following timing.

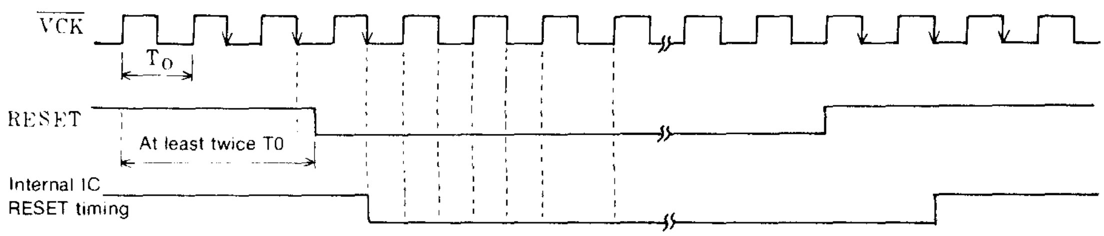



**Fig. 39 MSM5205 reset timing**

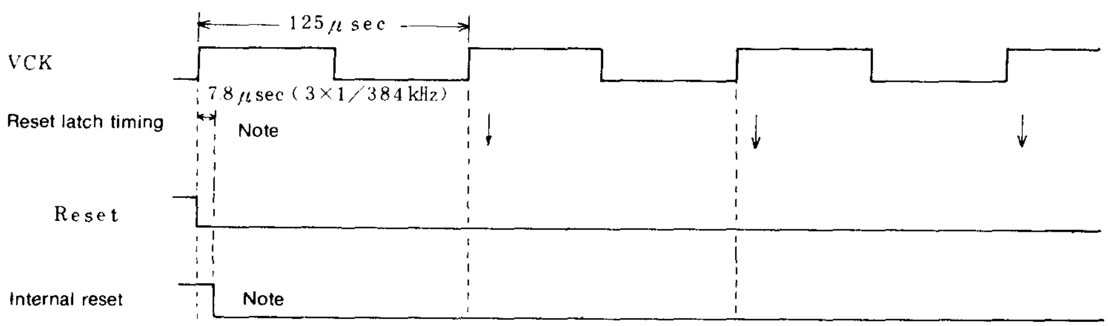



**Fig. 40 MSM5218 reset timing (8 kHz sampling example)**

Note: The reset signal is latched within the LSI by the reset latch timing. Analysis is commenced by switching the external reset signal from H to L before this timing. Switching is probably best achieved by the leading edge of the VCK signal.

**Q8** Why is a low-pass filter required in the voice synthesis output?

**A8** Low-pass filters (LPF) are designed to pass only those frequency components below a certain frequency when the input contains a number of different components. Since the voice synthesis output is obtained from a DA converter output, the output waveform is a stepwise waveform as indicated in Figure 41.

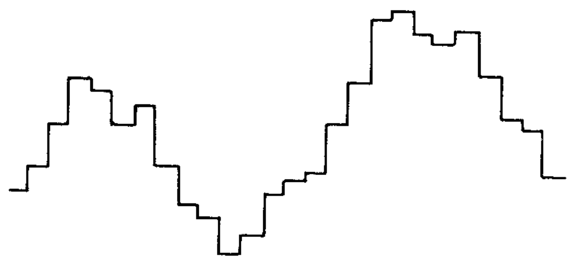


**Fig. 41 DA converter voice output waveform**

This waveform contains certain high frequency noise components, and since the sampling theorem states that sampling is not effective unless only frequency components below half the sampling frequency are obtained, the unwanted high frequency components are removed by the LPF. The LPF output waveform is as shown in Figure 42.

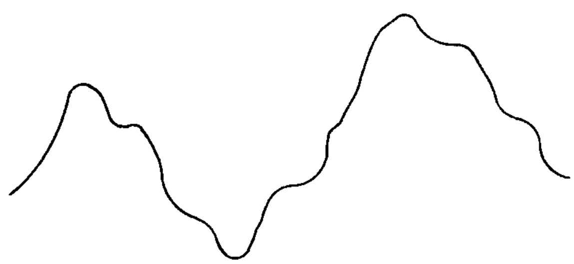


**Fig. 42 LPF output waveform**

**Q9** How is the LPF cut-off frequency best determined?

**A9** In an ideal LPF, all frequencies above a certain level are filtered out. In practice, however, frequencies are attenuated at a certain rate (slope) from a particular level. This slope is usually expressed in terms of dB per octave (dB/oct). For example, in a filter with a slope of -12 dB/oct, the output is attenuated by -12 dB (that is, reduced to 1/4) for each octave increase of the frequency (that is, each time the frequency is doubled).
The frequency where the attenuation commences is called the cut-off frequency (note). The best cut-off frequency for any situation will depend on the sampling frequency, the attenuation slope, and the frequency components in the source sound. The usual design criteria is the degree of attenuation (that is, how many dB) at half the sampling frequency ($f_{SAM}$).
A number of reference values are listed in Table 2. Since the optimum value as determined for the frequency components in the source sound varies considerably, ddit is best to decide on the basis of actual listening.

**Table 2 Relationship between filter attenuation characteristics, cut-off frequency, and sampling frequency**

| Filter attenuation characteristics (dB/oct) | Cut-off frequency | Gain at 0.5 $f_{SAM}$ (dB) | Diagram number |
|---|---|---|---|
| -12 | 0.3 $f_{SAM}$ | - 8.7 | ① |
| -18 | 0.33 $f_{SAM}$ | -11 | — |
| -24 | 0.35 $f_{SAM}$ | -12.5 | ② |
| -48 | 0.38 $f_{SAM}$ | -19 | — |

Although high frequency noise components are attenuated more efficiently when the attenuation characteristics are "sharper", a larger number of structural elements are required, making such a filter uneconomical. The filter frequency response when $f_{SAMPLE}$ is 8 kHz is outlined in Figure 43 for reference purposes.

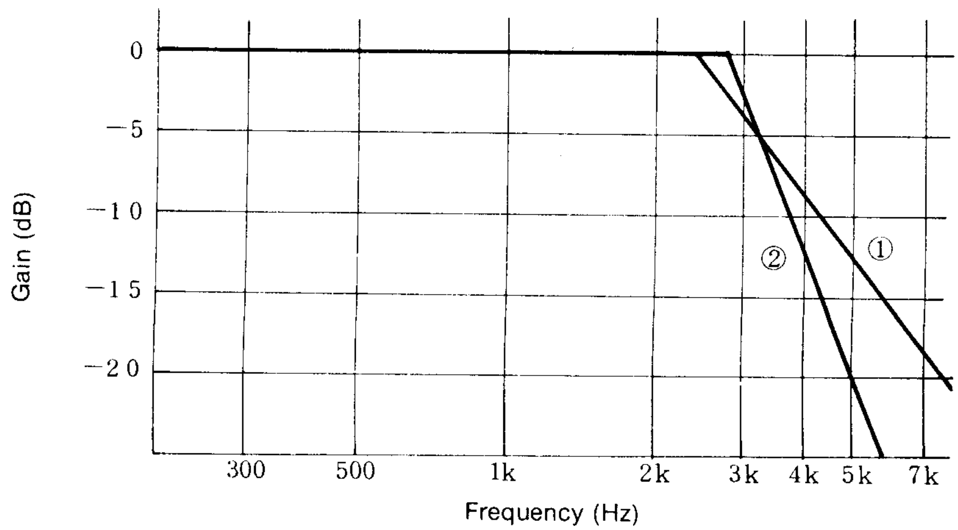



**Fig. 43 $f_{SAMPLE}$ 8 kHz filter frequency response**

Note: Strictly speaking, the cut-off frequency is not defined as the point where attenuation is commenced. In Butterworth type filters, the cut-off frequency is the "point -3 dB below the overall characteristics", while in Chebyshev type filters it is defined as the "point where the maximum ripple width within the set passband first appears". Where strict adherence to either definition is not required, the cut-off frequency can be assumed to be the point where attenuation is commenced.

**Q10** What is the configuration of low-cost filters designed for voice synthesis applications?

**A10** The most common types of active filters (filters with active elements) used are the Butterworth, Bessel, and Chebyshev filters. Each type is used for different purposes.

The Butterworth filter puts emphasis on the flatness of the passband and is less strict in terms of attenuation characteristics and transient response performance than the Bessel and Chebyshev filters.

In the LPF used for voice synthesis where the passband attenuation characteristics do not need to be perfectly flat, a Chebyshev filter capable of achieving sharp attenuation characteristics with a small number of component parts should be considered. Chebyshev filters can be designed with suitable ripple width and attenuation characteristics.

Note that if the frequency response of the speaker rolls off at the desired cut-off frequency, the LPF will not be required.

**Configuration**

A 1-pole RC LPF where no active elements are used is shown in Figure 44, and the corresponding transmission characteristics can be expressed in the following way.

$$F(j\omega) = \frac{\ell o}{\ell i} = \frac{1}{1 + i\left(\frac{\omega}{\omega_0}\right)} \quad \cdots\cdots\cdots\cdots \quad (1)$$

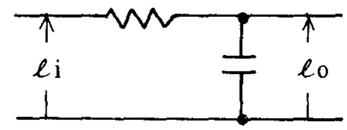



**Fig. 44 1-pole RC LPF**

where

$$\omega_0 = \frac{1}{CR} \quad (\omega_0 = 2\pi f_0)$$

fo: cut-off frequency

An F(jw) plot is shown in Figure 45.

The frequency response shown in this diagram is cut off at -6 dB/oct above frequency wo where $w_0$ is the -3 dB value.

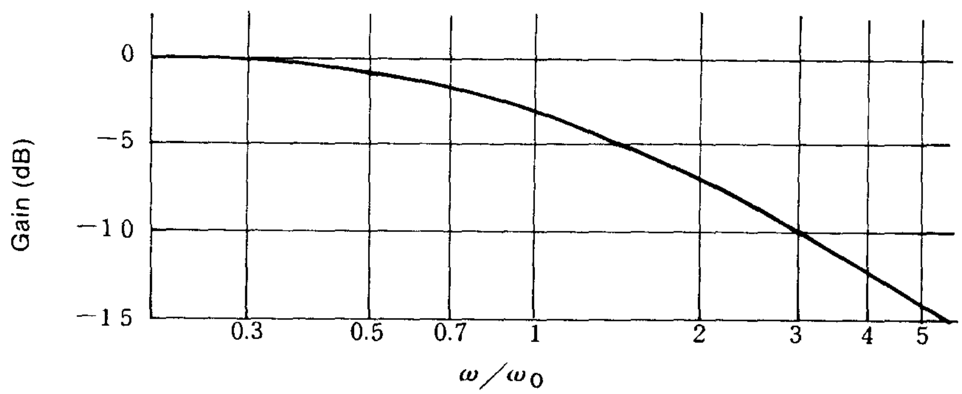



**Fig. 45 CR1 stage LPF frequency response**

The circuit configuration of a 2nd order Chebyshev filter is shown in Figure 46.

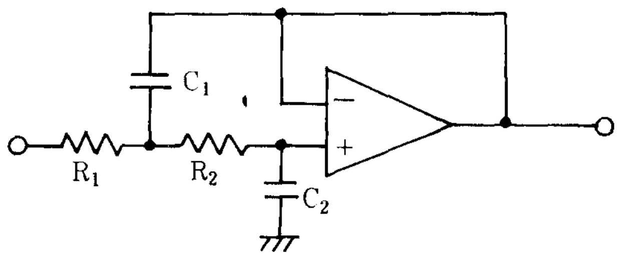



**Fig. 46 2nd order Chebyshev type filter**

The transmission characteristics for this circuit can be expressed in the following way.

$$F(j\omega) = \frac{\ell o}{\ell i} \quad \frac{1}{1 - \left(\frac{\omega}{\omega_0}\right)^2 + j\frac{1}{Q} \cdot \frac{\omega}{\omega_0}}$$

$$Q = \frac{\sqrt{R_1 C_1 R_2 C_2}}{C_2 (R_1 + R_2)} \qquad \omega_0 = \frac{1}{\sqrt{R_1 C_1 R_2 C_2}}$$

The active element used in this circuit is assumed to be an OP amplifier voltage follower with a gain of 1.

$$C_1 = \frac{2Q}{\omega_0 R} \qquad C_2 = \frac{1}{2Q\omega_0 R}$$

## Higher order Chebyshev filter design

Even numbered higher order filters such as 4th and 6th order filters can be divided into 2nd order elements. And odd numbered higher order filters such as 3rd and 5th order filters can be divided into 2nd order and 1st order (CR1 stage passive filter) elements.

A 4th order filter, for example, can be divided into two 2nd order element stages as shown in Figure 47. And by determining the fn and qn factors in each stage, the filter configuration can be easily achieved.

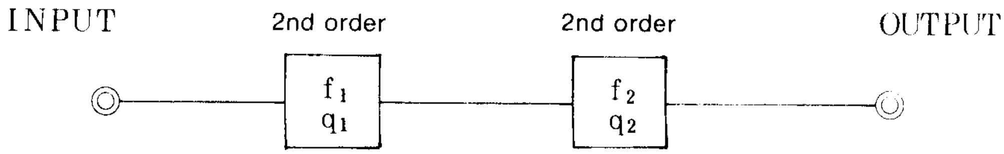



**Fig. 47 4th order LPF**

The attenuation characteristics of Chebyshev filters can be changed depending on the ripple tolerance within the passband. The fn and qn settings will also vary according to the tolerance.

A list of fn and qn values for Chebyshev filters is given in Table 3 below.

**Table 3 Chebyshev LPF and HPF fn and qn values**

| | Ripple = 0.1 dB | | Ripple = 0.2 dB | | Ripple = 0.25 dB | | Ripple = 0.3 dB | | Ripple = 0.5 dB | |
|---|---|---|---|---|---|---|---|---|---|---|
| | fn | qn | fn | qn | fn | qn | fn | qn | fn | qn |
| 2nd order | 1.8204497 | 0.7673593 | 1.5351966 | 0.7966418 | 1.4539722 | 0.8092536 | 1.3911667 | 0.8210811 | 1.231418 | 0.8637210 |
| 3rd order | 1.2999029<br>0.9694057 | 1.3490276<br>0.5* | 1.1889612<br>0.8146341 | 1.4595033<br>0.5* | 1.1569921<br>0.7672227 | 1.5080264<br>0.5* | 1.1321861<br>0.7292773 | 1.5524768<br>0.5* | 1.0688535<br>0.6254565 | 1.7061895<br>0.5* |
| 4th order | 1.1532699<br>0.7892557 | 2.1829303<br>0.6188010 | 1.0948338<br>0.7011094 | 2.4350125<br>0.6458968 | 1.0779389<br>0.6744223 | 2.5361100<br>0.6572494 | 1.0648159<br>0.6532428 | 2.6279020<br>0.6677803 | 1.0312704<br>0.5970024 | 2.9405542<br>0.7051102 |
| 5th order | 1.0931318<br>0.7974460<br>0.5389143 | 3.2820141<br>0.9145215<br>0.5* | 1.0570753<br>0.7472558<br>0.4614106 | 3.7068586<br>1.0009079<br>0.5* | 1.0466301<br>0.7324054<br>0.4369509 | 3.8756825<br>1.0359319<br>0.5* | 1.0385110<br>0.7207553<br>0.4171291 | 4.0283601<br>1.0678979<br>0.5* | 1.0177347<br>0.6904832<br>0.3623196 | 4.5449633<br>1.1778056<br>0.5* |
| 6th order | 1.0627261<br>0.8344903<br>0.5131875 | 4.6329012<br>1.3315707<br>0.5994600 | 1.0382299<br>0.8030621<br>0.4603216 | 5.2689021<br>1.4917187<br>0.6259511 | 1.0311242<br>0.7938542<br>0.4440628 | 5.5204164<br>1.5556533<br>0.6370268 | 1.0255981<br>0.7866630<br>0.4310754 | 5.7474076<br>1.6135959<br>0.6472924 | 1.0114459<br>0.7681212<br>0.3962290 | 6.5128456<br>1.8103772<br>0.6836390 |

The 0.5* in the qn column denotes CR1 stage 1st order

**Example**

The following example shows how a 5th order Chebyshev filter would actually be designed. A ripple tolerance of 0.5 dB is assumed. The circuit diagram is shown in Figure 48.

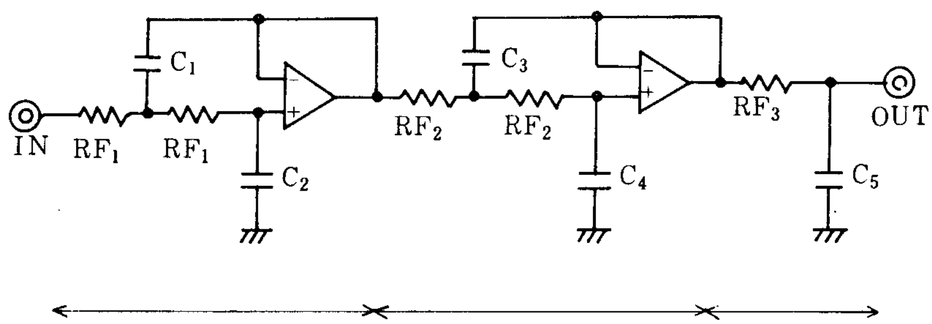



**Fig. 48 5th order Chebyshev LPF**

The fn and qn values have already been listed in Table 3. Where ripple 0.5 dB for a 5th order filter, the 1st stage 2nd order fn=1.0177347 and qn=4.5449633 the 2nd stage 2nd order fn=0.6904832 and qn=1.1778056 the 3rd stage 1st order fn=0.3623196 and qn=0.5. The constants can thus be calculated using these values. (begin-ind) A cut-off frequency of 2.8 kHz is selected. This value is then substituted in equations introduced earlier for calculating 1st and 2nd order constants.

1st stage 2nd order

Assume $R_{F1}$ = 51 kohm

$$f_0 = 2800 \times 1.0177347 \div 2850(\mathrm{Hz})$$

$$C_1 = \frac{2Q}{\omega_0 R_F} = \frac{2qn}{2\pi f_0 R_F} = \frac{2 \times 4.5449633}{2\pi \times 2850 \times 51 \times 10^3} = 9953(\mathrm{pF})$$

$$C_2 = \frac{1}{2Q\omega_0 R_F} = \frac{1}{2qn2\pi f_0 R_F} = \frac{1}{2 \times 4.5449633 \times 2\pi \times 2850 \times 51 \times 10^3}$$

$$= 120(\mathrm{pF})$$

2nd stage 2nd order

Assume $R_{F2}$ = 56 kohm

$$f_0 = 2800 \times 0.6904832 \div 1933(\mathrm{Hz})$$

$$C_3 = \frac{2Q}{\omega_0 R_F} = \frac{2 \times 1.1778056}{2\pi \times 1933 \times 56 \times 10^3} = 3463(\mathrm{pF})$$

$$C_4 = \frac{1}{2Q\omega_0 R_F} = \frac{1}{2 \times 1.1778056 \times 2\pi \times 1933 \times 56 \times 10^3} = 624(\mathrm{pF})$$

3rd stage 1st order

Assume $R_{F3}$ = 68 kohm

$$f_0 = 2800 \times 0.3623196 \div 1014(\mathrm{Hz})$$

$$C_5 = \frac{1}{\omega_0 R_F} = \frac{1}{2\pi f_0 R_{F3}} = \frac{1}{2\pi \times 1014 \times 68 \times 10^3} = 2308(\mathrm{pF})$$

In this case, the RF value can be changed to approach the value of a real capacitor. If the values of C1 and C2 are increased 1.5 times, RF1 is more or less divided by 1.5. By selecting values which approach capacitor values, the filter constants finally selected are as shown in Figure 49.

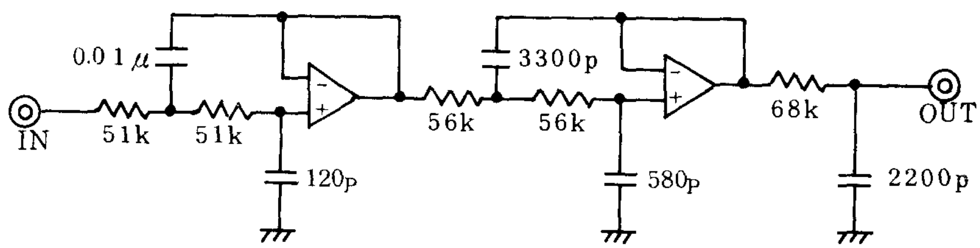



**Fig. 49 Designed filter (5th order)**

The filter characteristics for this filter can then be plotted. The transmission characteristics for the 1st and 2nd stage 2nd order filter elements can be expressed in the following way.

$$F(j\omega) = \frac{1}{1\ (\frac{\omega}{\omega_0})^2 + j\frac{1}{Q} \cdot \frac{\omega}{\omega_0}}$$

And the transmission characteristics for the 3rd stage 1st order filter element can be expressed by

$$F(j\omega) = \frac{1}{1 + j(\frac{\omega}{\omega_0})}$$

When, for example, w wo in the 1st stage filter,

$$F(j\omega_0) = \frac{1}{1 - (1)^2 + j\frac{1}{Q} \times 1} = -jQ = -4.5449633j$$

If the absolute input/output voltage ratio is expressed in dB,

$$20 \log 4.545 = 13.15 \text{ (dB)}$$

That is, the level is 13.15 dB when the frequency is 2850 Hz. The plot obtained inserting successive values in this equation is shown as a dashed line in Figure 50. And similar plots obtained for the 2nd and 3rd stages are shown as the dot-dash and double-dot-dash lines. The overall characteristics are shown by the full line.
line.

According to the overall characteristics, the use of values approaching actual capacitors gives a maximum ripple of 0.6 dB and a cut-off frequency of about 2.9 kHz.

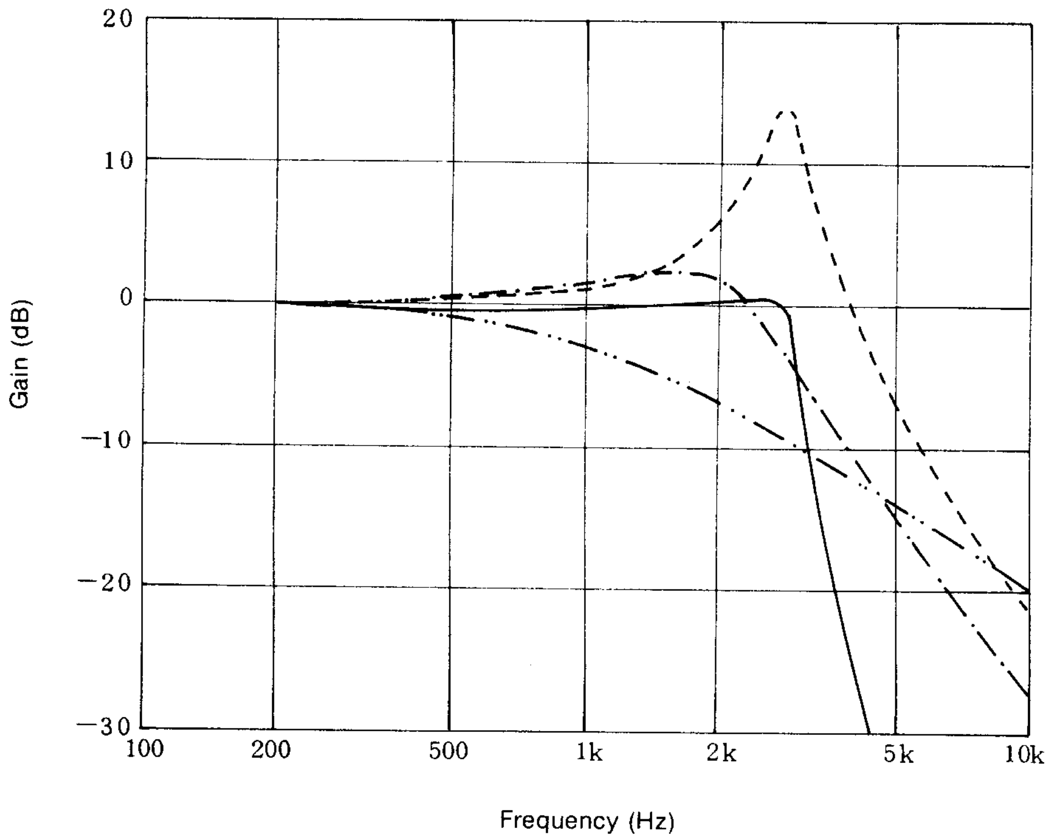



**Fig. 50 Frequency response of the designed filter (5th order)**

The constants and frequency response for a 3rd order Chebyshev filter designed in the same manner are shown in Figures 51 and 52.

The full line in Figure 52 represents the overall characteristics.

According to these overall characteristics, the use of values approaching actual capacitors gives a maximum ripple of 3 dB and a cut-off frequency of about 2.56 kHz.

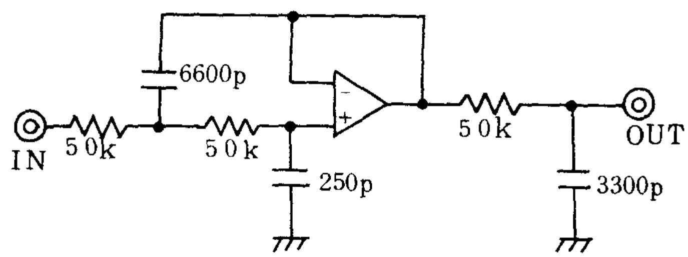



**Fig. 51 Designed filter (3rd order)**

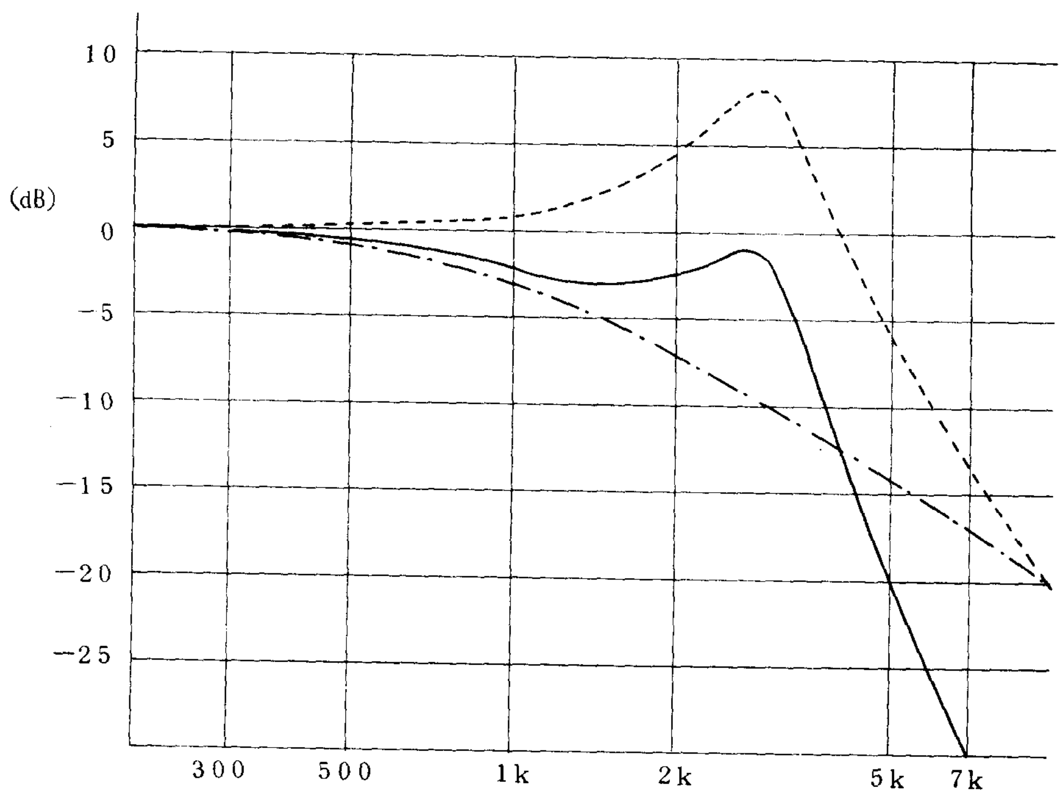



**Fig. 52 Designed filter frequency response (3rd order)**

**Q11** How should the filter be connected where the MSM6243 voice output pin output impedance is 60 kohm (typ)?

**A11** An output impedance of 60 kohm is equivalent to connecting 60 kohms to a 0 ohm output impedance pin (ideal constant voltage source).
If ambient temperature fluctuations and manufacturing tolerance spread are considered, this 60 kohm output impedance can vary from -50% to 100%, making filter constant settings impossible.
Hence, it is recommended that a voltage follower be inserted before the filter as shown in Figure 53 to vary the impedance.
(A voltage follower should also be used when MSM5205, MSM5218, MSM6212, MSM6258, and other devices are used.)

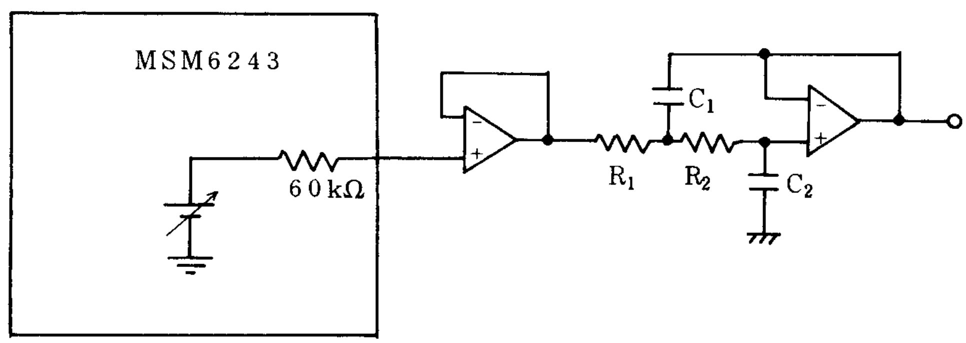



**Fig. 53 MSM6243 and filter interface**

**Q12** Can the MSM5205 voice output be connected directly to an audio amplifier?

**A12** **Yes. But note that the amplifier input impedance and allowable input amplitude must be considered.**

With a preamplifier stage input impedance of about 47 kohm, the normal input amplitude is 200 mVp-p. Since the MSM5205 voice output impedance is approximately 100 kohm and the output amplitude about 4 Vp-p ($V_{DD}$ = 5 V when unloaded), a load has to be connected to reduce the amplitude to about 1/20. Likewise with other devices, direct connection to an audio amplifier is enabled by connecting a load to match the output impedance. See Figure 54.

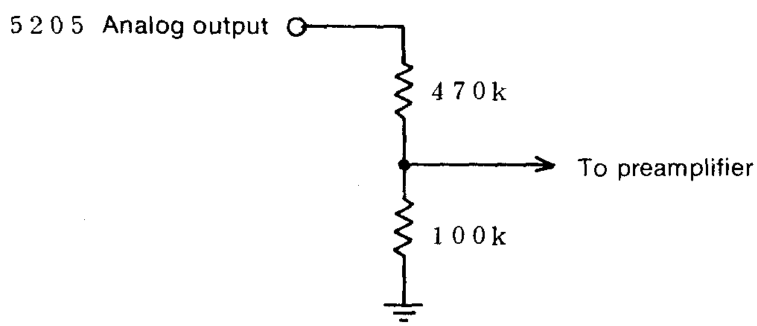



**Fig. 54 Audio amplifier connection**